RADC-TR-74-270
Final Report
October 1974

VOLUME I
APPENDICES

STANDARD PROCEDURES FOR AIR FORCE OPERATIONAL TEST AND EVALUATION

Braddock, Dunn & McDonald, Incorporated*

Distribution limited to US Gov't agencies only; test and evaluation; October 1974. Other requests for this document must be referred to RADC (IRAA), GAFB, NY 13441.

*with subcontractors RCA Government and Commercial Systems Missile and Surface Radar Division and the Xerox Corporation Data System Division

Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, New York 13441

Best
Available
Copy

VOLUME I
APPENDICES

STANDARD PROCEDURES FOR AIR FORCE OPERATIONAL TEST AND EVALUATION

D.E. Simon (RCA)
et al

Braddock, Dunn & McDonald, Incorporated*

Distribution limited to US Gov't agencies only; test and evaluation; October 1974. Other requests for this document must be referred to RADC (IRAA), GAFB, NY 13441.

*with subcontractors RCA Government and Commercial Systems Missile and Surface Radar Division and the Xerox Corporation Data System Division

Appendix A

DOCUMENTATION MODULES

1. THE REQUIRED OPERATIONAL CAPABILITY (ROC)

Broad Definition. A ROC is a formal document originated by a major command to identify and assess an operational deficiency or need to HQ USAF. ROCS are numbered consecutively by major commands during each calendar year and are submitted, after appropriate coordination, to HQ USAF/RDQLM for the initiation of responsive action.

Relationship to System Acquisition Cycle. The ROC serves as the initiating document, marking the beginning of the requirements process and the conceptual phase of the acquisition cycle.

Originator. Although the ROC (the formal statement of the operational deficiency or need) must be submitted by a major command and signed by a general officer or a colonel occupying a general officer position, any echelon of the Air Force or Department of Defense may identify the operational deficiency or need.

Purpose. The ROC presents the initial formal statement of the assessment of operational need to HQ USAF in order to establish a clear understanding of the need, urgency and importance. The need may result from the identification of a deficiency in the capability required to effectively counter a current or projected threat, or it may result from a technological breakthrough which offers a potential increase in operational effectiveness or efficiency.

Content. The ROC contains the following:

a. Command Heading

(1) Title (the identification of the ROC, limited to forty (40) characters and spaces to facilitate mechanized reporting.)

(2) ROC Number (the originating major command (e.g., TAC), followed by the consecutive number assignment of the ROC for the calendar year of submission (e.g., 5-74, the fifth ROC submitted by TAC in calendar year 1974)).

(3) Preparing Office - This entry will include office symbol and action officer's rank, name and telephone number.

(4) Date of release by the command section of the originating major command.

b. Primary Sections. In addition to the Command Heading, the ROC must consist of two (2) primary sections identified as follows:

(1) Section I. Deficiencies/Needs - This section contains a description of the operational deficiency or need and the assigned mission or task affected.

(2) Section II. Required Operational Capability - This section is a description of the required operational capability and the envisioned concept of operations including mission scenarios, operating environment, and deployment concepts. In addition, influences on the design solution such as operational, threat and survivability data, support concepts and constraints should be clearly stated.

The command authority line and signature of a general officer or colonel occupying a general officer position will follow Section II.

c. Discretionary Attachments. At the discretion of the originating major command, one or more attachments which include, but are not limited to the following may be submitted with the ROC. Each attachment will bear the ROC number and date in the upper right hand corner:

(1) Section III. Determinations of Deficiencies/Needs and the Required Operational Capability - The contents of this section provide an

explanation of how the stated deficiency or need was identified and what means are currently being employed to accomplish the mission or task. A brief description will also be provided concerning the effect of the deficiency or mission performance weighed against the current and future threat. Studies and analyses which support the ROC will be summarized and referenced in this section.

(2) Section IV. Solutions - A brief description of any and all solutions considered by the originating major command is appropriate for this section. Detailed technical specifications (which tend to restrict the technical approach to a design solution) are not desired. All proposed solutions should be listed in order of preference. The rationale for the order of preference should be included.

(3) Section V. Class V Modifications - Where a Class V Modification appears to be the best solution, Section V may be attached to the ROC. Section V will include the following:

(a) The evaluation and approval of the proposal for a Class V Modification by the command Configuration Control Board or equivalent.

(b) A description of the proposed modification.

(c) A description of the system or equipment in sufficient detail to permit ready identification of the item(s) to be modified or employed in the modification.

(d) A description of the development or qualification status, if known, of the equipment to be used in the modification.

(e) The command recommendation on testing, kit proofing, and prototype.

(f) Recommendations for level of accomplishment of the modification, i.e., field, depot, etc.

(g) Effect of the proposed modification on support equipment, spares, simulators, trainers, etc.

(h) Effect on manpower requirements.

(i) Training requirements.

(4) Section VI. Quantities Involved - This section permits a command estimate of the quantitative requirements with supporting data.

(5) Section VII. Special Comments - As deemed necessary by the originating major command, Section VII may be submitted as an attachment to the ROC to provide additional comments relative to the required capability and proposed technical solutions. Appropriate comments under this section may relate to such matters as maintainability, reliability, compatibility, security, crew comfort, and proposed Initial Operational Capability (IOC) dates, or any other known information that is appropriate.

Distribution of ROCs. Distribution by the originating major command will be made within seven (7) days of the date of the ROC as follows:

a. Action copies (25 each) to HQ USAF/RDQLM

b. Advance Information Copies (5 each) to:

(1) HQ USAF/RDR (if pertaining to Reconnaissance and Electronics Warfare)

(2) HQ USAF/RDS (if pertaining to Space)

c. Information Copies and Harmonization Copies. As directed by AFR 57-1.

Coordination. Coordination is required prior to the submission of the ROC by the major command. Further coordination is accomplished as a function of the validation of the ROC after receipt of the ROC by HQ USAF/RDQLM.

a. The pre-submission coordination will be affected between the major command initiating the ROC and:

(1) The using command to insure adequate visibility of user requirements.

(2) AFSC to ascertain technical feasibility.

(3) AFLC, as appropriate, for engineering responsibility related to the requirement.

(Note: Either AFSC or AFLC will review all draft ROCs, provide information to the major command on technologies which may be applied to resolving the operational deficiency or need, and assist the major command in the preparation of a ROC which meets the needs of the requirements process.)

b. Post submission coordination of the ROC, managed by HQ USAF/RDQLM, comprises the ROC validation effort. The ROC validation entails the following coordination:

(1) Review of the ROC and submission of comments to HQ USAF by implementing and participating commands.

(2) The establishment of a Program Element Monitor (PEM) for the ROC within HQ USAF/RDQ.

(3) The review of the ROC by the Requirements Review Group (RRG) within HQ USAF, with the decision to either validate (approve) the ROC or to disapprove the ROC.

(4) The actions within HQ USAF/X00 to prioritize the ROC in relation to other ROCs. The assignment of funds for the conceptual phase studies and analyses is based on this prioritization.

(5) Coordination, as appropriate, with other Services, other departments and agencies and allies of the United States.

ROC Inputs and Outputs. There are various influences which produce the decision by a major command to generate a ROC. The ROC also has primary effect on the acquisition of systems and equipment and on the determination of test objectives for both Development Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E).

a. Some of the influences which may provide an input to the ROC are:

(1) The identification of a new or changing threat on the basis of intelligence information.

(2) A study or analysis of data at the major command or subordinate command level which reveals deficiencies or needs in mission capability resulting from inadequate resources, tactics, or other man-made or natural constraints.

b. The ROC influences the following:

(1) The identification and propagation of a stated deficiency or need by the originating major command.

(2) The evaluation of the alternative means to satisfy the deficiency or need, and the ranking of the deficiency or need in relation to other requirements.

(3) The generation of formal HQ USAF directives in the conceptual phase to authorize, direct and control the necessary actions required to identify and move the desired solution into the mainstream of the acquisition process.

(4) The establishment of the basic test objectives in the test and evaluation process, i.e., the satisfaction of the deficiencies or needs set forth in the ROC.

References. The following references were used in the preparation of this text:

a. AFR 57-1; "Policies, Responsibilities, and Procedures for Obtaining New and Improved Operational Capabilities."

b. HQ USAF/RDQLM unclassified message 102213Z October 1973.

2. THE COMBAT REQUIRED OPERATIONAL CAPABILITY (COMBAT ROC)

Broad Definition. The Combat ROC is a documented mission - essential requirement for an operational capability which directly affects the ability of USAF combat forces to conduct combat operations. Combat ROC's are assigned consecutive numbers during each calendar year prefixed by the unit's abbreviation, i.e., 8AF 5-74, 8AF 6-74, etc., and suffixed, as appropriate, by the foreign force abbreviation in parentheses should the Combat ROC be a Military Assistance Service Funded (MASF) requirement. An example of the latter is 8AF 7-74 (ROKAF). The Combat ROC is simultaneously disseminated, upon submission by the originating Combat Command, to key addressees to facilitate concurrent staffing by all agencies without direction of and prior to approval by HQ USAF.

Relationship to System Acquisition Process. As in the case of the ROC, the Combat ROC serves as the initiating document of the requirements process, marking the beginning of the conceptual phase of the acquisition cycle.

Originator. Any USAF Combat Commander may originate a Combat ROC.

Purpose. The Combat ROC expresses the operational needs of USAF combat forces engaged in or imminently threatened with combat operations against hostile forces.

Content. The Combat ROC normally contains only the basic information required by key addressees for action and decision. The format and content guidance provided herein and in AFR 57-1 for the ROC is applicable also to the Combat ROC. For example, the inclusion of Section I (Deficiencies/Needs) and Section II (Required Operational Capability) in the Combat ROC is mandatory. Should the originator desire to include information in addition to that prescribed for Sections I and II, such additional information should be in the format and under the appropriate heading prescribed for the ROC, Sections III through VII.

Distribution of Combat ROCs. The initial submission of a Combat ROC may be by either message or letter; however, message submission is encouraged for all communications pertaining to Combat ROCs. When a Combat ROC is transmitted by

message, the originator will reproduce the message on DD Form 173 and mail a copy of the reproduction to each message addressee not later than 2 workdays following message transmission. The initial distribution of Combat ROCs by the Combat Commander is as follows:

a. Action copy addressees:

Combat Area Commander
HQ USAF/RDQ
Appropriate System Manager/Item Manager
AFSC/DA
ASD/RWT
ATC/XPQ
AFSC/XOT
TAC/DR
*HQ USAF/RDR (reconnaissance and electronic warfare items)
*USAF SS (communications security equipment)
*AFCS/XRQ (communications-electronics or Air Traffic Control)
*MAC/DOQ (airlift, search and rescue, recovery, weather reconnaissance, armament recording, documentation photography and television)
*AFLC/MMA (Class V Modifications)

b. Information copy addressees:

HQ USAF/LGY
HQ USAF/XOO

Coordination and actions. Upon receipt of the Combat ROC by action addressees, coordination and staffing actions are taken within prescribed time limits in

*Additional action addressees for specific involvement as indicated.

compliance with AFR 57-1. The "immediate" and "follow-up" actions required of addressees are as follows:

a. Combat Area Commander

(Actions upon receipt of Combat ROC)

(1) Review, verify need for, prioritize with respect to all other active (not satisfied or withdrawn) Combat ROCs from his area.

(2) Forward message to all addressees and originating combat commander within 5 work days stating tentative approval or disapproval.

(Follow-up actions)

(3) Within 20 work days following the receipt of the Combat ROC, forward message to all addressees and originating combat commander stating final approval or disapproval.

(4) Conduct quarterly review of each active Combat ROC submitted from his area to determine if requirement is still valid. Prepare priority listing of all active (including MASF Class V Modification) Combat ROCs. Submit copies of priority list to HQ USAF/RDQ/RDP/RDR/XOO to arrive by first day of 2nd, 5th, 8th, and 11th calendar months.

b. HQ AFSC

(Actions upon receipt of Combat ROC)

(1) Review and staffing. Commence tentative explorations for technical solutions to problem.

(2) Upon receipt of tentative approval by Combat Area Commander, commence preparation of a best planning estimate (BPE) of technical solution to problem.

(Follow-up actions)

(3) Within 20 workdays following receipt of Combat Area Commander's tentative approval, forward BPE by message to USAF/RDQ (and/RDR if reconnaissance or electronic warfare items are involved). Information copies to originating Combat Commander and appropriate Combat Area Commander, HQ USAF/LGY,/XOO, AFSC/SDA, TAC/DR, AFLC/XOT, ATC/XPQ, MAC/DOQ.

(4) Provide information on Class V Modification Combat ROCs to AFLC as needed.

c. HQ TAC

(Actions upon receipt of Combat ROC)

(1) Review; staff coordination.

(2) Upon receipt of Combat Area Commander's tentative approval, begin assessment of impact of requirement, the proposed tactical concept, and the operational utility of the end-product.

(Follow-up actions)

(3) Within 10 workdays following receipt of the BPE, forward by message the TAC evaluation of the proposed solution and appropriate comments to ASD/RWT, and to the originator and all addressee recipients of the BPE.

(4) Provide operational guidance on the effect of the proposed solution on world wide tactical air capabilities.

d. HQ AFLC

(Actions upon receipt of Combat ROC)

(1) Review; staff coordination. Identify AFLC involvement.

(2) Where Class V Modification is involved, and upon receipt of Combat Area Commander's tentative approval, direct the preparation of a

Class V Modification best planning estimate (BPE) by the appropriate System Manager/Item Manager.

(3) Coordinate with ATC/XPQ and ASD/RW (tactical warfare) or the appropriate SPO.

(Follow-Up Actions)

(4) Within 20 workdays following receipt of Combat area Commander's tentative approval, forward BPE by message to: HQ USAF/RDQ (and/RDR if reconnaissance and electronic warfare items are involved) with information copies to originating Combat Commander and appropriate Combat Area Commander; HQ USAF/LGY, /XOO; HQ AFSC/SDA; HQ TAC/DR; HQ ATC/XPQ; HQ MAC/DOQ.

e. Combat Commander/Combat Area Commander. Within 10 workdays following the receipt of the message BPE from either AFSC or AFLC, submit comments and recommendations concerning the BPE to HQ USAF/RDQ (and to/RDR if appropriate) with information copies to HQ USAF/LGY and/XOO.

f. HQ USAF

(Actions taken upon receipt of Combat ROC)

(1) Commence staff review and validation effort.

(2) Upon receipt of the Combat Area Commanders tentative approval, or as promptly as practicable, approve or disapprove the Combat ROC.

(Follow-up actions)

(3) Within 20 workdays following the receipt of comments and recommendations resulting from the review of the BPE by the Combat Commander and Combat Area Commander, issue implementing instructions in Program Management Directive (PMD) form.

(4) Advise all concerned of actions taken and status on each Combat ROC.

Combat ROC Inputs and Outputs. The influences which dictate the initiation of and which result from a Combat ROC are:

a. Input Influences.

(1) The involvement of a USAF command in combat operations or immediately threatened with combat operations, and

(2) The identification of a mission-essential need for an operational capability which directly affects the commander's ability to conduct combat operations.

b. The output influences are:

(1) The action, within specific due dates, at Combat Area Command Headquarters for tentative and final approval or disapproval of the requirement.

(2) The concurrent staffing and coordination by all concerned agencies within specific due dates.

(3) The generation of the proposed technical solution, (the best planning estimate (BPE)) and the review and approval of the BPE within a specific due date.

(4) The issuance of implementing program instructions (PMD) within specific due dates to provide the required combat capability.

Reference. The reference used in the preparation of this text was AFR 57-1; "Policies, Responsibilities, and Procedures for Obtaining New and Improved Operational Capabilities".

3. THE QUICK REACTION CAPABILITY (QRC)

Broad Definition. The Quick Reaction Capability (QRC) is a management and acquisition process used within the Air Force to provide rapid response to urgent operational deficiencies or needs in electronic warfare, reconnaissance, and intelligence. Normally, all written communications and program management directives on all QRC matters are sent by electrical message form.

Relationship to System Acquisition Cycle. The QRC procedures are divorced from the normal acquisition process and phases within the process. For QRC procedures, the submission of a QRC request by a major command constitutes the beginning of the process which, in a normal acquisition cycle, would be the conceptual phase.

Originator. Any major command may submit a request for a QRC.

Purpose. The purpose of QRC is to minimize the time interval between the emergence of a new threat and the establishment of an operational and support capability in the areas of electronic warfare, reconnaissance or intelligence to cope with the threat. Except for unusually urgent requirements approved by HQ USAF, QRC actions may be applied only to programs involving small numbers of systems or equipment.

Content. Requests for a QRC will include the following:

a. Description of the required capability which clearly states and justifies the need for urgent solution.

b. Description of anticipated operational concept, environment, and use.

c. Quantity required, logistic considerations, including basis for issue and estimated flying/operating hour program.

d. Required delivery schedule.

e. System(s) to which applicable.

Distribution of Request for QRC. The major command submitting the request for QRC will submit copies of the request to action and information addressees listed as follows:

a. Action copy addresses:

HQ USAF/RDRE

b. Information copy addresses:

HQ USAF/XOO,/RDQ,/INY,/LGY

AFSC	PACAF
AFLC	ADC
SAC	AAC
TAC	ATC
USAFE	AFSS

Coordination and Action. Upon receipt of the request for QRC, coordination and staffing actions are taken within prescribed time limits by both action and information addresses in compliance with AFR 57-5 and as follows:

a. Other commands with similar needs:

(1) Within 10 workdays following receipt of request for QRC, submit requirements to HQ USAF/RDRE stating any unique considerations. Provide information copies to all recipients of the originating major command request.

(2) A negative reply is required.

b. HQ USAF/RDRE

(1) Within 20 workdays following receipt of the originating major command's request, either disapprove the request for QRC or direct AFSC/AFLC, as appropriate, to conduct an evaluation and feasibility study.

(2) Appraise originating major command and other interested parties of action taken.

c. HQ AFSC/AFLC:

(1) Within 40 workdays following receipt of HQ USAF/RDRE direction, provide findings of evaluation and feasibility study to HQUSAF/RDRE with information copies to the originating major command and other commands supporting the QRC.

(2) Conduct an evaluation and feasibility study of the request for QRC to address the following areas:

(a) Total or partial satisfaction of the requirement

(b) Alternative solutions

(c) Recommended solutions

(d) Estimated schedules and costs (complete project)

(e) Cost breakdown: including follow-on production, development, production, test, logistics, support.

(f) Cost estimates (any associated system modification)

(g) Requirements, if any, for unusual resources, assistance or priorities.

(h) Test and test support aircraft required in development or test.

d. HQ USAF/RDRE: Within 20 workdays following receipt of findings from AFSC/AFLC, either approve the development of the QRC in accordance with findings, direct further evaluation, or disapprove the program. In the event of approval, HQ USAF/RDRE takes further action to:

(1) Assign QRC program designator and title

(2) Provide program milestones and equipment performance parameters

(3) Designate implementing and participating command(s) and assign responsibility

(4) Identify approved funding

(5) Provide for test and test support aircraft requirements cited in the AFSC/AFLC findings of the evaluation and feasibility study.

(6) When installation of a QRC production item requires modification to a related system, issue modification program management directive concurrent with QRC program approval.

c. The implementing command (AFSC or AFLC):

(1) Exercises maximum authority and responsibility delegated by HQ USAF to manage QRC program within funding, schedule, and performance constraints.

(2) Expedites development, engineering, test, and logistics to accomplish QRC program objectives within specified time frames.

(3) Advises HQ USAF of any inhibiting policy or procedure which requires waiver or change.

(4) Appoints the QRC program manager.

(5) Prepares required program documentation for support of QRC.

Related QRC Responsibilities.

a. AFSC

(1) Utilizing monthly QRC management reporting system (RCS: SYS R-29), provide description, status and milestones on all active QRC programs, including AFLC-managed QRC programs, with distribution as directed by HQ USAF.

(2) Engineering assistance and system/item management control as required.

(3) Assistance to AFLC in preparation of evaluations and feasibility studies.

(4) Provide schedules and priorities to accomplish testing of QRC items.

b. AFLC

(1) Inputs to AFSC for monthly QRC report

(2) Management of QRC Logistic Support Plan

(3) Inventory manager for all QRC production items, selected QRC development items and specialized support equipment.

(4) Operational engineering support for QRC items.

(5) Assistance to AFSC in QRC item installation when requested.

c. ATC

(1) Monitor equipment maintenance/operator procedures in DT&E/IOT&E testing of QRC equipment and verify adequacy of procedures.

(2) Act as consultant to the implementing command on technical order verification and on support equipment.

d. Other Major Commands

(1) Establish an OPR within the command for QRC programs with procedures for expediting QRC matters.

(2) Appoint a primary and alternate QRC officer; furnish names, office symbols and telephone numbers to HQ USAF, AFLC and AFSC.

(3) Provide a project officer in TDY status when requested by AFSC/AFLC to assist in QRC test and evaluation.

(4) Perform specific QRC program tasks as assigned.

QRC Inputs and Outputs.

a. The influences which dictate the initiation of a request for QRC are:

(1) The identification of an existing or postulated threat involving electronic warfare, reconnaissance and intelligence activities.

(2) The estimation of enemy advantage posed by the threat.

(3) The need for effective, prompt countermeasures to the threat.

b. The output influences of the QRC are:

(1) Concurrent coordination and actions, within specific due dates by action agencies, by the approval authority, and by other commands having operational interests similar to the requesting major command.

(2) The dynamic evolvement of a responsive management and acquisition program for the engineering development of a solution, and the pro duction, test, evaluation and deployment of QRC items.

(3) The establishment of the basic test and evaluation objectives of the QRC item.

Reference. AFR 57-5; "Quick Reaction Capability".

4. THE DEVELOPMENT CONCEPT PAPER (DCP)

Broad Definition. The Development Concept Paper (DCP) is an OSD (Office of the Secretary of Defense) document which serves in the management of a major system development program to:

a. Present the OSD Staff management position and rationale supporting a determination of need for a major system, including a plan for the acquisition of the system and,

b. State the decisions of the Secretary of Defense (SEC DEF) including the decisions to approve or disapprove

   (1) The initiation of the validation phase (the program decision)

   (2) The transition of the program from validation phase to full-scale development phase (the ratification decision)

   (3) The transition of the program from full-scale development phase to production phase (the production decision)

Relationship to System Acquisition Cycle. The DCP is initially drafted during the latter period of the conceptual phase. With the supporting documentation of studies and analyses, it forms the program advocacy package presented for DSARC I review and to the SEC DEF for the first key decision point (the program decision). Subsequently, the DCP is updated as necessary to accurately contain the summary information of the program, the decision review thresholds, and the management issues which requires the SEC DEF decision. In addition to its role in the review and decision process which permits the initiation of the validation phase (DSARC 1), the DCP is also the principal document for program review at DSARCs 2 and 3 and the related SEC DEF decisions points (ratification decision and production decision).

The decision review thresholds mentioned above are governing limits which, if exceeded or when expected to be exceeded, require a SEC DEF program review. In such instances, an updated DCP must be presented as the document vehicle for the review.

Originator. The Director of Defense Research and Engineering (DDR&E) (or the Assistant Secretary of Defense (Telecommunications) for his programs) within Office, Secretary of Defense has the basic responsibility for the coordination of inputs for the DCP and for the submittal for the DSARC review and the SEC DEF decision. The Air Force is committed to full and complete support to OSD in such matters and has established a counterpart office within HQ USAF for DCPs, identified as the Assistant for Research and Development Programming (DCS/R&D). This office serves as the Air Staff central point of contact on all matters pertaining to DCPs related to Air Force programs. As a function of providing full and complete support, the Program Element Monitor/Program Officer (PEM) within the Air Staff Directorate having program responsibility is tasked with the responsibility for the preparation of the draft DCP and updating and revising the DCP, coordinating such actions within the Air Staff, AFSC and the Program Manager as necessary, and also with his OSD counterpart, the OSD Project Officer for the major system acquisition program.

Purpose. The purpose of the DCP is to provide summary documentation of a major defense system acquisition program in order to support the DSARC reviews and the SEC DEF decision making process throughout the acquisition process.

Content. The DCP consists of a cover sheet, main body, and a signature page. The information content of the DCP is prescribed in Attachment 3, HQ USAF Operating Instruction (HOI) 11-16 "Responsibilities, Functions and Procedures Pertaining to Development Concept Papers (DCPs)". The content is confined to the essential summary information required for the SEC DEF program management decision. The DCP cover sheet identifies the particular issue of the DCP, the preparation date, the responsible Service component and field agency, the responsible ODDR&E unit and specific major system program identifiers. The main body of the DCP addresses the major subject headings listed below in succinct, summary information appropriate for the SEC DEF level:

a. The Management Issue

b. The Reasons for Having the Program ("The Problem")

c. Ways of Solving "The Problem" (Possible "Solutions")

d. Anticipated Resource Needs of the Program

e. Anticipated Schedule for the Program

f. Risks the Program Faces

g. Program Difficulties

h. Evaluation

i. Decision Options

j. Recommendations

Distribution of DCPs. The distribution of DCPs is accomplished by the PEM within the Air Staff Directorate having program responsibility and reflects the coordination and information needs of both OSD and the Air Force. The development of the DCP, initiated by the DCP outline, is an iterative process progressing through the development of the draft DCP, the OSD "for comment" draft, the OSD "for coordination" draft, and the approved DCP with different distribution needs for each. In general, distribution is made as follows:

a. Distribution of the DCP outline: (first iteration)

(1) OSD Project Officer and within OSD, as appropriate.

(2) Air Staff PEM and within HQ USAF, as appropriate.

(3) AFSC and the Program Manager.

b. Distribution of the draft DCP: Through Air (second iteration) Force, as necessary for comments coordination, proposed changes. Following the Air Force coordination, the draft DCP is revised by the Air Staff PEM to incorporate appropriate changes/recommendations and the document then becomes the OSD "for comment" draft. (third iteration)

c. Distribution of the OSD "for coordination" draft DCP:

(1) OSD

(2) AFSC

(3) Program Manager

(4) Throughout DOD as directed by OSD for comments, coordination and proposed changes.

Following the completion of DOD coordination, OSD provides proposed change material to the Air Staff PEM, who revises the DCP to include required changes. The document becomes the "OSD for coordination" draft. (fourth iteration)

d. Distribution of the OSD "for coordination" draft DCP:

(1) OSD

(2) DSARC principals

(3) ASD (T) or ASD (I), as appropriate

(4) Chairman, Joint Chiefs of Staff or his designee

(5) Deputy Director (T&E) DDR&E

(6) Secretary of the Air Force

(7) AFSC

(8) Program Manager

e. The approved DCP - Following the DSARC, the decision of the SEC DEF is set forth in a decision memorandum to SAF. The PEM in the Directorate of the Air Staff having program responsibility updates the DCP to reflect the SEC DEF decision and within 10 workdays following the decision distributes copies of the approved DCP to: (fifth iteration)

(1) DSARC prinicpals

(2) SAF, and within Air Staff as appropriate

(3) AFSC

(4) Program Manager

(5) Others, as appropriate

Coordination and Actions. A high degree of coordination is accomplished in each evolutionary phase of the DCP, commencing with the coordination involved in the agreement between OSD and HQ USAF on DCP outline and proceeding through the development of the draft DCP, the OSD "for comment" draft, the OSD "for coordination" draft, and the approved DCP. For each phase of DCP evolution, coordination and actions as listed below are required:

a. Coordination incident to formalization of the DCP outline:

(1) (Air Staff PEM/OSD Project Officer). Informal coordination on a day-to-day basis to include problems such as required studies and analyses to support the program, force issues, status of technology, management organizations and funding.

(2) (Air Staff PEM) - Coordinate with Air Staff, AFSC and Program Manager on proposed DCP outline.

(3) (Air Staff PEM/OSD Project Officer) - Arrive at an agreement on the DCP outline.

(4) (Air Staff PEM) - Preparation of the DCP outline. Provide copies to Assistant for Research and Development Programming, DCS/R&D for distribution within Air Force. Provide copies to OSD Project Officer.

(5) (SAF/RD) - Formally submit briefing outline to OSD for comment.

(6) (Air Staff PEM) - Obtain written approval for DCP outline from OSD Project Officer.

b. Coordination for the preparation of the draft DCP:

(1) (Air Staff PEM/Program Manager/AFSC) - Compile the draft DCP. Allow a minimum of six weeks for preparation of AFSC draft inputs.

(2) (Air Staff PEM) - Distribute draft DCP to implementing/supporting/using commands, Air Staff, Program Manager for coordination.

(3) (Air Staff PEM) - Consolidate Air Force comments. Provide copies to AFSC and Program Manager.

(4) (CSAF) - Submit draft DCP to OSD at pre-DSARC planning meeting (draft DCP becomes OSD "for comments" draft).

c. Coordination of OSD "for comment" draft DCP

(1) (OSD Project Officer) Distribute OSD "for comment" draft throughout DOD for comments.

(2) (OSD Project Officer) - Provide OSD review and comments to Air Staff PEM.

(3) (Air Staff PEM) - Consolidate OSD comments in second draft DCP.

(4) (Air Staff PEM) - Submit second draft to DSARC principals at least 10 days prior to the scheduled DSARC. The second draft DCP becomes the OSD "for coordination" draft.

d. Coordination of OSD "for coordination" draft DCP:

(1) (OSD Project Officer). Coordinates draft DCP with DSARC principals (and ASD (T)) or ASDI as appropriate), the Chairman of the Joint Chiefs of Staff or his designee, the Deputy Director (T&E), DDR&E, and the Secretary of the Air Force.

(2) (OSD Project Officer). Provide OSD comments resulting from OSD coordination to HQ USAF and DSARC chairman not less than 5 workdays prior to the scheduled DSARC.

(3) (Air Staff PEM) Forward copies of OSD comments to AFSC and Program Manager for DSARC preparation.

e. Coordination of approved DCP:

(1) Following the DSARC, the decision of the Secretary of Defense is set forth in a decision memorandum from Secretary of Defense to Secretary of the Air Force.

(2) (Air Staff PEM) - Updates the DCP, incorporating the Sec Def decision. Within 10 workdays after the SEC DEF decision is made, distributes copies of the approved, updated DCP to DSARC principals, AFSC, Program Manager, and others as appropriate.

DCP Inputs and Outputs. The influences which dictate the initiation of the DCPs and which result from the DCP are set forth below.

a. Input influences:

(1) The DOD documentation requirements of a major system acquisition program, in which the DCP is required as the OSD decision document reflecting program objectives, accomplishment, future plans, and decisions.

(2) The requirement for advocacy documentation to gain SEC DEF approval to advance the program acquisition process of a major system.

(3) To appraise SEC DEF of a breach in pre-established decision thresholds in a major system acquisition programs and to facilitate the program review and SEC DEF decision required in such instances.

b. The influences of the outline and draft versions of the DCP are:

(1) The selection and conduct of specific studies and analyses required in support of the DCP.

(2) The formulation of the program objectives, program issues, plans, performance parameters, system alternatives, logistics, acquisition

strategy, milestone decision points, thresholds in cost, time and performance, and the critical questions and issues to be addressed by the test and evaluation program.

(3) The program presentation to the DSARC, the recommendations of the DSARC to SEC DEF and the SEC DEF decision at key program decision points to either continue in the acquisition of a system, or to cancel, alter the direction of, or alter the pace of the program.

c. The influences of the approved DCP are:

(1) The approved DCP serves as the contract between OSD and HQ USAF and the authorization for Air Force to proceed with the conduct of specifically bounded activities involving the acquisition process for a major system.

(2) The approved DCP influences the actions prescribed in subsequent Program Management Directives (PMD) used by HQ USAF to implement the program.

(3) The approved DCP, and the supporting documentation concerning test and evaluation (i.e., the draft Test Plan, test program reports) addresses critical questions and issues and related test objectives. In this respect, the DCP influences the plan for test and evaluation contained in the Test Objectives Annex (TOA) of the Program Management Plan (PMP).

References.

a. DODI 5000.2; "The Decision Coordinating Paper and Defense Systems Acquisition Review Council." (Draft)

b. AFR 800-2; "Program Management"

c. HQ USAF HOI 11-16; "Responsibilities, Functions and Procedures Pertaining to Development Concept Papers (DCPs)".

d. HQ USAF HOI 800-1; "DCP/DSARC Preparation".

e. AFSCP 800-3; "A Guide for Program Management".

5. THE PROGRAM MEMORANDUM (PM)

Broad Definition. The Program Memorandum (PM) is sometimes called a "mini-DCP (Development Concept Paper)." Like the DCP, the Program Memorandum is an OSD (Office, Secretary of Defense) document which serves in the management of a system development program to present the OSD Staff management position and rationale supporting the need for a system, including a plan for the acquisition of the system. Like the DCP, the PM serves also to document program guidelines, thresholds, and the decisions of the review authority concerning the development program.

The PM differs from the DCP in that the PM is applied to less-than-major system development programs while the DCP is required for major system programs. The use of the PM is reserved for development programs which, while not meeting the definition of major system development programs, are nonetheless sufficiently important to merit the retention of program review and decision authority within OSD. The development program to which a Program Memorandum is applied is not subject to DSARC review or the SECDEF level of program decision. Normally the decision level for a program under a Program Memorandum is one level below the DCP decision level (e.g., at the Director DDR&E level vice the SEC DEF level), with program decision authority remaining within OSD.

Relationship to System Acqusiition Cycle. The relationship of the PM to the system acquisition process is the same as is presented herein for the DCP. The PM is originated late in the conceptual phase.

Originator. Same as for DCP.

Purposes. The purpose of the PM is to provide summary documentation of a significantly important (but less-than-major) defense system acquisition program in order to support the program briefings and decision making process in OSD throughout the acquisition process of the system. The approved PM, as in the case of an approved DCP, constitutes approval authority to implement the program described by and within thresholds specified in the PM.

Content. The format and content of the PM is the same as that prescribed for the DCP.

Distribution of PMs. The distribution of PMs is the same as accorded DCP, except that review/decision authority should be substituted for DSARC Principals wherever appearing.

Coordination and Actions. Same as for the DCP, except that appropriate review/decision authority should be substituted for DSARC Principals and SEC DEF where-ever appearing.

PM Inputs and Outputs. Except for the specific references to DSARC and the SEC DEF decision process, the influences affecting the PM and those influences exerted by the PM on the system acquisition process are identical to those of the DCP.

References. Only general references are available at this writing concerning the Program Memorandum. Both AFLCM 800-1; "Acquisition Management Concept" and AFSCP 800-3; "A Guide For Program Management" contain fragmentary information on the purpose and composition of the PM. Additional information on the use and format of the PM and the decision-level authority for PM-related systems acquisition programs was obtained through contact with responsible offices within the Air Staff.

6. THE PROGRAM MANAGEMENT DIRECTIVE (PMD)

Broad Definition. The Program Management Directive (PMD) is the official HQ USAF management directive used during the entire system/equipment acquisition cycle to provide direction to the implementing and participating commands and satisfy requirements for the documentation of the program.

Relationship to System Acquisition Cycle. The PMD is a living document which is initiated in the conceptual phase and is updated as necessary to reflect the approved management direction for the program. The PMD serves throughout all phases of the acquisition cycle. Prior to the program decision (DSARC 1), the PMD exists as a formal HQ USAF action directive.

Originator. The Program Officer within the assigned Office of Primary Responsibility (OPR) is the originator of the PMD. During the conceptual phase, the OPR for the PMD is HQ USAF/RDQ. Commencing with the program decision at DSARC 1 the OPR functions are transferred from RDQ to the directorate or special Air Staff office exercising primary interest responsibilities. Any Air Staff directorate or special Air Staff office may issue a PMD for areas in which they are clearly identified as the Air Staff OPR.

Purpose. The PMD is used throughout the acquisition life cycle to state requirements, request studies, and initiate, modify, approve, change, transition or terminate programs.

Content. The content of the PMD is the prerogative of the designated OPR, subject to appropriate coordination. It provides the specific program direction or information required at that point in time. While as brief as possible, the contents of the PMD may range from a paragraph to several pages and may address only a selected program area or all program areas. For example, a PMD written by the Program Officer/Program Element Monitor (PEM) in the early stages of the conceptual phase to direct the conduct of a study to provide the technology roadmap for a design solution may be limited to only the task and the required inputs to the task. In comparison, the sequentially numbered issue of that PMD which follows the program decision

at DSARC 1 would normally provide complete program direction and would address all program areas. Thus, the content of the PMD varies, reflecting not only the phase and point in phase of a program, but also the complexity, cost, size, importance and organizational interrelationships of the program. Generally, the contents of the PMD may be broadly categorized as (1) contractual or guidance information, and (2) action directive, with elements of each as follows:

a. Contractual or guidance:

(1) Contract management information between HQ USAF and the implementing and participating commands

(2) Program decisions and agreements

(3) Approval/disapproval actions by review authorities

(4) Guidance and definition

(5) Program monitoring and review thresholds

(6) Resources and funding information

b. Action directive:

(1) Conduct studies and analyses of stated operational deficiencies

(2) Conduct preliminary design studies

(3) Development of solutions or technological alternatives

(4) Preparation of program advocacy documentation

(5) Management of retrofit configuration changes

(6) Authorize and allocate, as practicable, test aircraft to the implementing command

(7) Initiate, modify, change, transition or terminate actions.

Subject Headings. There are specific candidate subject headings appropriate for the PMD as set forth in the PMD example format in Attachment 1 to HQ USAF HDI 800-2. The format is provided as a guide for writing a PMD and only the headings appropriate for the purpose of the PMD will be used. These headings and a summary of paragraph content consisting of information, actions required or guidance, as applicable, are:

a. Paragraph 1. SPECIFIC PURPOSE: (Mandatory for inclusion in all PMDs)

(1) Provides the reason for the PMD

(2) Identifies what the PMD is intended to do

(3) Identifies the project

(4) States the task required, who will perform the task (by name and address of each action addressee) and when the task will be performed

b. Paragraph 2. PROGRAM SUMMARY:

(1) Summary of program information relative to program origin

(2) Priority and objectives

(3) References

(4) Force structure information

(5) Management concept of phases in program life cycle

c. Paragraph 3. INTELLIGENCE/THREAT ESTIMATE: Scope, level of detail and procedures for updating the intelligence/threat estimate

d. Paragraph 4. PROGRAM MANAGEMENT:

(1) Complete description of required action

(2) Action required of each action addressee

(3) Schedule milestones and due dates

(4) Assign and limit authority

(5) Reporting requirements

(6) Resources and support

(7) Precautionary guidance

(8) Contract and source selection guidance

(9) Life cycle cost estimates

(10) Instructions as appropriate for transition or termination

e. Paragraph 5. SYSTEMS ENGINEERING:

(1) Scientific and engineering application to the total systems engineering effort

(2) Performance parameters

(3) Configuration and systems design

f. Paragraph 6. TEST AND EVALUATION:

(1) Critical questions and issues which must be evaluated

(2) Extent of participation by implementing and participating commands

(3) Test Objective Annex (TOA) to the PMD

(4) Independent evaluations by implementing, operating, and supporting commands

(5) Requirements to coordinate and comply with AFR 80-14

(6) Authorization and allocation of test aircraft and associated manpower, as appropriate

(7) Determination of test support resources requirements

g. Paragraph 7. COMMUNICATIONS/ELECTRONICS: Special requirements in the operational system or test support program

h. Paragraph 8. OPERATIONS: Current operations plan or changes thereto

i. Paragraph 9. CIVIL ENGINEERING:

(1) Use of present facilities

(2) Funding new requirements

j. Paragraph 10. LOGISTICS: Unique or unusual logistics guidance or information (other than direction for inclusion of the Integrated Logistics Support flow provided by AFR 800-8).

k. Paragraph 11. MANPOWER:

(1) Estimated manpower requirements related to the system for each command and functional area.

(2) Coordination of changes and requirements with using commands

(3) Statement of whether manpower requirements will be met from existing resources or additional manpower is to be programmed by HQ USAF.

l. Paragraph 12. DATA AUTOMATION: Guidance for selection, acquisition, development, utilization and management.

m. Paragraph 13. PERSONNEL TRAINING:

(1) Training required in response to the system

(2) Time phasing for training

(3) Equipment support and procurement

n. Paragraph 14. SECURITY:

(1) Address requirement for Security Classification Guide and review/ approval by the Assistant Secretary of the Air Force, Research and Development

(2) Specific directions for classification requirements in administration, design, performance, operations, vulnerability, deficiency, etc.

(3) Special security considerations, including access by foreign nationals.

o. Paragraph 15. PUBLIC RELEASE:

(1) Details of public release restrictions or requirements.

(2) Release of information to the public by contractors.

p. Paragraph 16. OTHER REQUIRED: The PMD may include guidance as required other than that specified in paragraphs 1 through 15.

Distribution of PMDs. A PMD may be issued in electrical message form if immediate notification is necessary. It should normally be issued within 30 days after receipt of program approval or change action from higher authority. The PMD will contain a distribution list prepared by the Program Officer. The distribution list will normally prescribe the following:

a. Action copies.

(1) 15 each to HQ AFSC or HQ AFLC if identified as implementing command

(2) 15 each to all other implementing commands

(3) 5 each to each participating command.

b. Information copies. As specified by the Program Officer, utilizing the listing contained in HQ USAF HOI 800-2, Attachment 4. Appropriate additions thereto may be made at the discretion of the Program Officer.

Coordination. The Program Officer is responsible for accomplishing all required coordination prior to the submission of the PMD for approval. Prior to the approval of the PMD at the directorate level or by higher authority, the official signing the

PMD will insure that the Program Officer has effected coordination with and provided a copy to all interested Air Staff officers, to include:

a. The appropriate Secretarial counterpart offices

b. HQ USAF/RDM for document control functions and adequacy of technical content and format.

c. HQ USAF/LGX, for PMDs involving the development of and funding for weapon systems, subsystems or related equipment.

d. The original PMD coordinating Air Staff office (for all follow-on PMDs)

e. HQ USAF/LGP, to ensure the inclusion in the PMD of relevant procurement or contractural instructions or appropriate guidance.

f. HQ USAF/RDPQ, for any program involving test aircraft.

g. The appropriate Air Staff focal point(s), to ensure that critical direction is included. Focal points and the number of coordination copies required by each are identified and listed in HQ USAF HOI 800-2, Attachment 2.

PMD Inputs and Outputs.

a. Inputs - Influences which result in the generation of a PMD by the Program Officer are related to the purposes of the PMD, set forth in paragraph 4 herein. Thus, the need to state requirements, request studies, and to initiate, modify, approve, change, transition or terminate programs in the acquisition process provides the inputs to the PMD.

b. Outputs - The PMD output influences the following:

(1) All activities of the Air Force in the acquisition process, commencing with the initial directive.

(2) The scope and content of the Program Management Plan (PMP)

(3) The authority and responsibility of the Program Manager.

(4) The program advocacy as presented at the program decision point (DSARC 1)

(5) The test and evaluation process for DT&E and IOT&E, including objectives, participants, resources, funding, and independent evaluations required.

References.

a. HQ USAF HOI 800-2 dated 1 Mar 73; subject: "Program Management Direction".

b. AFR 57-1 dated 17 Aug 71; subject: "Policies, Responsibilities, and Procedures For Obtaining New and Improved Operational Capabilities".

c. AFR 800-2 dated 16 Mar 72; subject: "Program Management".

d. AFLCM 800-1 dated 29 Dec 72; subject: "Program Management".

e. AFSCP 800-3 dated 14 May 71; subject: "A Guide For Program Management".

7. THE TEST OBJECTIVES ANNEX TO THE PROGRAM MANAGEMENT DIRECTIVE

Broad Definition. The Test objectives Annex (TOA) to the Program Management Directive (PMD) is a HQ USAF management directive written in the style and level of detail of the PMD and which supplements the PMD to provide the baseline for the test program of a major system from which independent evaluations by operating, supporting, and development commands are required. The TOA is prepared in response to specific direction of HQ USAF for certain major programs.

Relationship to System Acquisition Cycle. The TOA is issued in the validation phase of the acquisition cycle and is updated with the PMD prior to each major program decision milestone supported by DCP preparation or revision such as the Defense Systems Acquisition Review Council (DSARC), or similar review by HQ USAF.

Originator. The TOA is drafted by the system program cadre office or Program Office within AFSC with operational test inputs provided by AFTEC, and is forwarded to the Program Officer/Program Element Monitor within HQ USAF for coordination, approval by the directorate having system responsibility, or higher authority as appropriate, and subsequent inclusion in the PMD.

Purpose. As an annex to the PMD, the purpose of the TOA is linked to the broader purpose of the PMD, which serves from the beginning of the validation phase throughout the remainder of the acquisition cycle as the official HQ USAF management document, providing direction to the implementing command and through the Program Office to the contractor, to the operating and supporting commands, and to all other commands and agencies involved in the development of the system. The TOA accomplishes such functions for the specific area of system testing. It provides the baseline for the preparation of subsequent system test and evaluation documentation, such as the test section of the Program Management Plan, the Program Test Plan and the Test Capabilities Annex, and the supporting Test Plans and test procedures including Test Plans involving the use of certain test ranges and facilities.

Certain major programs in which development and operational risks are high and which involve the expenditure of a great quantity of resources demand a higher level

of management control and supervision. In such instances, the TOA elevates the control authority for the test and evaluation baseline from the Program Office level to the HQ USAF level, ensuring management visibility in test and evaluation of such significant major acquisitions program.

Content. There is no prescribed format for the TOA. The scope and depth of coverage in the TOA will normally be determined by the HQ USAF directorate responsible for system development in coordination with the Program Office. The content will depend upon size, complexity, time span of the program resources required, program decision needs at all levels within the DOD, and the urgency of the test program. The content will depend also upon the phase of system development in relation to the total acquisition cycle. This is due to the fact that critical questions and issues are not static in nature. As system development evolves, certain critical questions and issues are resolved and others are surfaced, requiring the updating of the TOA and changes of content, as appropriate, to reflect the resolution of certain critical questions and issues and the identification, association of objectives, and direction for resolution of newly identified critical questions and issues. There are certain content requirements, generally stated, for the TOA which are listed as follows:

a. The complete, current listing of critical questions and issues compiled by HQ AFSC from the developer (including the contractor), and the supporter.

b. The critical questions and issues of an operational nature provided by AFTEC.

c. The critical question and issues introduced by the DSARC review which require address by test and evaluation prior to the next DSARC review.

d. A discussion of each critical question and issue to provide background information and the rationale for inclusion as a critical question or issue.

e. The general test and evaluation approach to be used in the resolution of critical questions and issues.

(NOTE: The DT&E inputs to the TOA are provided by the implementing command. The OT&E inputs are provided by AFTEC.)

f. The identification of specific test objectives which respond to each critical question and issue.

g. Tasking of commands for T&E and time phasing for goals.

h. Test support resources and requirements.

i. General reporting requirements.

Distribution of the TOA. The Program Officer/Program Element Monitor within the HQ USAF directorate having program responsibility prescribes the distribution for the TOA, which normally follows the distribution accorded the PMD.

Coordination. As early as practicable in the definition of the program, the development organization convenes a test conference attended by representatives of the developing organization, the system contractor, AFTEC, appropriate test organizations, operating and supporting commands, and other headquarters, as appropriate. The subject of this initial test conference is the formulation of test objectives. The drafting of the TOA is a coordinated effort among these participants, among whom the representative of the development organization guides and directs the conference activity.

Following the coordination noted above, the TOA is coordinated within HQ AFSC and AFTEC and upon approval at those levels is submitted as the draft TOA to the Program Officer/Program Element Monitor within the responsible directorate in the Air Staff for coordination within HQ USAF.

TOA Inputs and Outputs. The influences which impact the content, size and format of the TOA, and, following its promulgation, the test and evaluation activities and concerns influenced by the TOA are as listed below:

a. Inputs to the TOA:

(1) Critical questions and issues identified by the contractor and the implementing, supporting and using commands and AFTEC.

(2) Critical questions and issues resulting from the DSARC reviews.

(3) Other risk and cost factors and the urgency of the program.

(4) Constraints imposed by availability of resources and state-of-the-art test technology.

(5) Data requirements for determining a level of confidence for system development.

(6) The identification of alternatives and the selection of objectives for the test program.

b. Output of the TOA. The TOA influences the following:

(1) Test planning, execution and reporting as set forth in the Test Section of the Program Management Plan (PMP) and in supporting test and evaluation planning documents prepared by AFTEC, the implementing and operating commands, and the supporting command.

(2) Management decisions concerning the development and deployment of the system, based on the relative importance of each critical question and issue, the selection of test objectives intended to resolve such matters, and the attainment of test objectives.

References. As of this writing, there is little reference material available concerning the TOA. The TOA is addressed briefly in paragraph 11.6 of AFSC Supplement 1 to AFR 80-14 titled "Test and Evaluation" and is mentioned in paragraph 6 "Test and Evaluation" of Attachment 1 to HQ USAF HOI 800-2 titled "Program Management Direction." Much of the content of the discussion of the TOA included herein was obtained from contacts with various responsible offices within HQ USAF.

8. THE PROGRAM MANAGEMENT PLAN (PMP)

Broad Definition. The Program Management Plan (PMP) is the principal management baseline document for implementing and planning the system development program. It is a thorough, coordinated and well-documented plan for managing acquisition programs. The PMP is developed by the Program Manager (PM) to present in greater scope and detail, and as integrated time-phased tasks and resources requirements, the tasks initially levied by the Program Management Directive (PMD) from HQ USAF and command supplements thereto. Unless otherwise directed in the PMD, the Program Manager approves the PMP and no approval by higher headquarters is required. The document contains only that information deemed necessary by the PM for the needs of the program. defines and identifies the participation and support responsibilities of participating organizations, and is directive upon the participating organizations. The PMP is updated by and at the discretion of the PM.

Relationship to System Acquisition Cycle. The preparation and promulgation of the PMP normally occurs early in the validation phase following the promulgation of the PMD and the approval of the program by the SEC DEF at the program decision milestone (DSARC 1). The PMP is updated by the PM at any point in the acquisition cycle in order to realign the PMP with new or changing program requirements as set forth in later PMD's and command supplements, and to affect changes in the integrated time-phased tasks and resources requirements.

Originator. The Program Manager (PM) is the originator of the PMP.

Purpose. The purpose of the PMP is twofold:

a. The PMP serves as the principal management baseline document during the validation, full-scale development, production and deployment phases of the acquisition cycle and to the point of transition of program responsibilities to AFLC.

b. The PMP serves as a directive in the levying of tasks, responsibilities, support functions and resources requirements on participating commands.

Content. The guidance for content and format of the PMP is contained in Attachment 4, AFSCP 800-3; "A Guide For Program Management," which is recommended as a source of valuable detail. There are several rules governing the content of the PMP, listed briefly as follows:

a. Include only the minimum essential information for outlining the overall management plan.

b. Limit detail to the implementing and planning requirements of the Program Office, HQ USAF, HQ ASFC and participating organizations.

c. Normally, the PMP need only reference in appropriate sections the supporting documents (such as Test Plans or Civil Engineering Master Plan) containing details of implementing and planning functions.

d. The amount of information included within a prescribed section may vary for each program depending upon factors such as system or equipment size, complexity, cost and the degree of national urgency attached to the program.

e. PMPs may be issued incrementally, depending on program urgency and the availability of input material. In such instances, the foreword will identify the document as an increment of the PMP and will include a proposed date for the next increment.

Subject Headings. There are also prescribed section and subject headings for the PMP as listed below. Additional sections may be included at the discretion of the Program Manager; however, each prescribed section should be addressed. If any of the subject headings listed below are not applicable to a program, an explanation should be included. The following is a list of sections, by subject content, presented for information and guidance. Specific instructions and format applicable to a program should be issued by the PM.

a. Section 1 - Program Summary and Authorization:

(1) Brief description of the system/equipment and rationale for selection, with applicable references to the DCP.

(2) Summary of the research and development and technology involved.

(3) The overall acquisition management concept.

(4) Summary of (or reference to) the latest PMD, AFSC Form 56, etc. that concerns program parameters, resources, or otherwise governs actions of participating organizations.

(5) Identify program priority, precedence rating, importance category as applicable.

b. Section 2 - Intelligence: (Provided to the Program Office by the intermediate command intelligence office)

(1) Summary of the threat baseline against which the program was initially approved for development, updated as changes occur. (Not to include information classified above SECRET. Reference documents which contain higher level classified intelligence information, as appropriate.)

(2) References to detailed intelligence/threat documentation related to the program.

(3) Description of the character and detail of intelligence/threat information continually required by the Program Office and participating organizations during acquisition to ensure that the system remains responsive to the threat.

c. Section 3 - Program Management:

(1) Overall management concept and approach (in greater detail than in Section 1).

(2) The establishing of an integrated management information system, responsive to Program Office and contractor needs.

(3) The PMP (Section 3) contains major subsections as appropriate identified as:

(a) Technical Performance - (continuous assessment of program accomplishments versus stated requirements)

(b) Schedules

Master Program/Overall Milestone Schedule
Production/Delivery Schedules
Facilities and Site Activation Schedules
Task Schedules
Training Schedules

(c) Interrelationships (organizational)

(d) Reporting Requirements

(e) Financial

(f) Procurement

(g) Production

(h) Contractor Data

(i) Turnover and Transition

(j) Risk Analysis

(k) Information

(l) Miscellaneous

d. Section 4 - System Engineering and Configuration Management: Provides the management concept for system engineering and configuration management.

e. Section 5 - Test and Evaluation:

(1) Provides the management concept for test and evaluation.

(2) Identifies all participating organizations.

(3) Identifies specific test objectives, (including critical issues and questions to be resolved by the test program), reporting requirements and procedures.

(4) Includes test plans developed in coordination with AFTEC and supporting commands for Development Test and Evaluation (DT&E) and Initial Operational Test and Evaluation (IOT&E).

(5) Provides for participation in Follow-on Operational Test and Evaluation as appropriate, and for engineering support through initial system deployment to an operational theater or base.

(6) Includes plans for AFTEC and supporting command participation in DT&E and other activities as mutually determined.

(7) Includes source of test support (government and contractor). The location of such support should be provided when determined, and, also, the approximate start/stop dates for the required support.

f. Section 6 - Communications/Electronics: Includes requirements for communications/electronics for Program Manager/ Program Office, test support, and the operational system.

g. Section 7 - Operations:

(1) Provides an expansion of the operational concept of the system under development, prepared with the assistance of the operating command.

(2) Supporting studies, analyses and documents should be cross-referenced.

(3) Subjects addressed within this section, as appropriate, are mission, limitations, deployment/operational plan and dates, command and control, readiness (including availability and reliability), the OT&E plan, unit maintenance, supply and safety, meteorological/environment, electronic warfare, organizational structure, transportation, personnel/manpower, training, facilities, special weapons, penetration aids, related training and operation readiness training, electromagnetic compatibility/electromagnetic environment/site surveys, and life support.

h. Section 8 - Civil Engineering:

(1) Includes a master plan where appropriate (or reference thereto) which outlines the proposed site development for total facility requirements, including each installation or sub-installation.

(2) Categorizes facility requirements as technical support real property (TSRP) or nontechnical support real property (NSRP).

(3) Assigns responsibility, establishes procedures, identifies commitments, programs, and resources for the necessary programming, design, construction, maintenance, acceptance and transfer (to the operating command) of real property requirements.

i. Section 9 - Logistics: (Prepared with inputs from and with assistance of the responsible logistics organization and other participating agencies.)

(1) Provides a comprehensive description of the tailored logistics concepts for the program.

(2) Includes considerations supporting integrated logistics applicable to the system/equipment planning, design, development, test demonstration and operational processes.

(3) Includes logistics program planning aspects concerning other elements of the PMP supporting reliability, maintainability, and transportability.

(4) Includes aspects of test equipment, supply support, transportation, packaging and handling, and technical data at all levels of logistic support.

(5) This section is coordinated with the Integrated Logistic Support Plan (ILSP) and its related management information system to ensure progress, status visibility, and overall life cycle logistics support.

j. Section 10 - Manpower and Organization:

(1) Provides a description of the organization of the Programs Office and the relationships and roles of other Air Force and government agencies involved in the acquisition.

(2) References any formal agreements with participating organization.

(3) Provides manpower requirements based on inputs from and with assistance of the Program Office, the operating command, other participating organizations, and the system/equipment contractors.

(4) States manpower requirements derived from operation and maintenance concepts and design parameters. Projects manpower requirements through the system/equipment life cycle.

(5) Includes organization charts and brief functional statements for the operating command units to which the system/equipment manpower will be allocated.

k. Section 11 - Personnel Training: (Prepared utilizing inputs from Air Training Command, operating commands, and other participating organizations). Provides summary of personnel training requirements to meet system/equipment tests and operational support activities, cross-referenced to other sections, as necessary, to reflect total PMP requirements in: requirements for trained personnel, types, location and key dates for individual training

courses, major items of required training equipment and schedules for activation; and initial and replacement training requirements by fiscal quarters, projected for 5 years, if applicable.

l. Section 12 - Security:

(1) Provides security support information and guidance concerning classification guidance, public release of information, release of information to foreign nationals, personnel security clearances, and industrial security.

(2) Provides system security information and guidance related to the design of the system/equipment, and the operational considerations concerning security.

m. Section 13 - Application of Directives: Contains the complete baseline list of directives, arranged in sections from 1 through 12 in consonance with the construction of the PMP, to provide the identity of directives which require an action by the Program Officer. For example, security directives which require an action of the Program Officer are listed within this section and are referenced to Section 12 (Security) of the PMP.

Coordination. As the singular program management baseline document for use by all participating agencies and higher-level decision authorities, the PMP requires a high degree of coordination between the Program Office and Air Force commands. Much of the content of the PMP is prepared by appropriate commands such as AFLC (Section 9. Logistics), the using command (Section 7. Operations), or Air Training Command (Section 11. Personnel Training), requiring the coordination of the Program Office and the responsible commands. Further coordination is required in the integration of time-phased tasks and resources requirements contained in the PMP. The degree of coordination involved in producing an effective, viable PMP depends upon the size of the program, the complexity of the program and the PMP required for the program, the number of agencies participating in the acquisition program and the

interfaces which evolve and are identified in the planning, compilation and implementation of the PMP. Coordination is normally accomplished between the Program Office and participating agencies through the representatives of participating agencies assigned to the Program Office and through the exercise of command/agency focal points established to effect coordination with the Program Office.

PMP Inputs and Outputs.

a. Inputs - The management plan requirements, as defined and outlined in Attachment 5, AFSCP 800-3; "A Guide For Program Management" determine the input requirements of the PMP. Such inputs must be in consonance with the Development Concept Paper (DCP), the Program Management Directive (PMD), and with command supplements to the PMD for the program.

b. Outputs - The PMP influences the following:

(1) The identification of tasks, responsibilities and resources and the identification and scheduling of integrated time-phased actions required of participants in the acquisition program.

(2) The degree of program acceptance at the key decision milestones in the acquisition process.

(3) All activities involved in the validation, development, production and deployment phases of the acquisition cycle.

(4) The planning and conduct of DT&E and IOT&E, the establishment of test objectives which respond to critical questions and issues, and the reporting of test results.

Reference. AFSCP 800-3 dated 14 May 1971, subject: "A Guide For Program Management."

9. THE TEST DIRECTIVE FOR OT&E

Broad Definition. The Headquarters Air Force Test Directive is a formal document, normally in the form of an official letter or electrically transmitted message which authorizes, directs, and provides the basic information necessary for the planning, execution, support and reporting of an OT&E program. The term test directive includes the specific documents in the evolution of the test directive which are:

a. The Initial Test Directive - A HQ USAF directive that initiates, for planning purposes, an IOT&E project. Sufficient authority and information is conveyed in this interim directive to form an initial test team and to develop the test design and Test Plan. The Initial Test Directive may be the PMD or a document identified specifically as the Initial Test Directive which is used in lieu of or in conjunction with a PMD to initiate test planning.

b. The Draft Test Directive. A document produced by AFTEC for submission to HQ USAF for review, approval, and upon approval, publication and dissemination as the HQ USAF Test Directive. The draft Test Directive is an expansion of the Initial Test Directive (or the PMD which served as the Initial Test Directive).

c. HQ USAF Test Directive. A final Test Directive which authorizes and directs the planning, execution, and reporting of an OT&E program. It also is a tasking document, requiring major commands to provide the required resources for an OT&E project assigned to AFTEC. The principal use of the Test Directive is to outline the purpose of the test, establish schedules, task participants, provide direction and guidance, and convey specific OT&E test direction that was not fully defined at the PMD issue date.

Relationship to System Acquisition Cycle. The relationship of the Test Directive to phases of the acquisition cycle is generally as follows:

a. For IOT&E

(1) The Initial Test Directive, which may be either a PMD or a letter or message-type document used for planning purposes, is promulgated as

early as is practicable in the acquisition process. The basis for the Initial Test Directive is the identification of critical questions and issues and the early inputs provided by AFTEC concerning OT&E requirements for the system, including the test schedule, resource estimate and draft test design.

(2) The Draft Test Directive is prepared by AFTEC subsequent to the issuance of the Initial Test Directive and is associated with either the validation phase or the early full-scale development phase.

(3) The HQ USAF (final) Test Directive is issued in sufficient time to permit detailed IOT&E planning and the conduct of IOT&E prior to the production decision (DSARC III). These considerations dictate that the document be issued in the validation phase or early in the full-scale development phase, and in any case, as soon as practicable following the coordination and approval of the Draft Test Directive at the HQ USAF level.

b. For FOT&E. The Test Directive for FOT&E is normally produced in the production and deployment phase. In some instances during an acquisition process a requirement may be foreseen for the conduct of FOT&E in the early production and deployment phase. In such instances an Initial Test Directive for FOT&E may be issued prior to the production decision in the full-scale development phase.

Originator. The AF/XOOW Project Officer, under the direction of the Deputy Director for Operational Test and Evaluation, Director for Operations, has the overall responsibility for the preparation, coordination, publication and distribution of USAF Test Directives for OT&E. In practice, the Project Officer prepares the Initial Test Directive utilizing inputs provided by AFTEC. For the HQ USAF (final) Test Directive, the Project Office coordinates the Draft Test Directive (prepared by AFTEC) among the Air Staff, performs the Air Staff functions for approval advocacy, and publishes and distributes the document after approval.

Purpose. The Test Directive identifies and tasks participants, identifies resources, provides authority, direction, guidance, essential command interface responsibilities, the rationale for, and baseline information concerning the planning, conduct, and reporting of OT&E The Test Directive is the authority and basis for implementing documents (the Project Order and the Project Plan) prepared by the major command tasked with conducting the OT&E.

Content. The level of detail contained in the Test Directive shall be appropriate for the needs of the specific program. There is no firmly prescribed format; however, a format and basic paragraph content is provided for use as a guide. This format and guidance is set forth in Attachment 1 to AF/XOOWD Deputy Director for Operational Test and Evaluation Operating Instructions (DD01) No. 10-12; subject "Test Directives", dated 5 April 1973 which contains the following:

Para. #	Heading	Content
1.	Title/Project Number	a. Nickname (approved by AF/XOOWB) or descriptive words for title. b. Project number (e.g. AFOTE/SAC/74/08) identifies test as a HQUSAF OT&E project to be conducted by the Strategic Air Command and is the eighth OT&E Test Directive issued by HQUSAF in FY 1974.
2.	Background	a. Provides historical basis for the initiation of the Test Directive. b. Includes significant facts, events, dates which led to the test requirement.
3.	Purpose	Provides overall purpose or requirement of the test, for example: operational deficiency, need for modification, new requirement for test data, documentation to support a solution, proposal or decision, etc.

Para. #	Heading	Content
4.	Concept of Operations	a. Describes overall intent of the test.
		b. Provides additional clarification of the purpose of the test.
		c. Identifies test and relationship of test to other testing.
		d. Indicates number and type of test articles involved.
		e. Describes overall tactics and techniques to be evaluated.
		f. Maintenance and logistic support concept.
		g. Expected test duration
		h. Special support personnel and other resources required.
5.	Test Objectives	a. Includes questions/issues to be addressed.
		b. States and defines specific test objectives.
		c. Considers needs, requirements and special interest of the user as well as expected hostile environment.
		d. Ensures that test objectives are in compliance with AFR 80-14, DODD 5000.3, and AF-HDI 800-2.

Para. #	Heading	Content
6.	Description of Test Items	a. Provides test items descriptions, including unique features that set it apart from other similar items.
		b. For aircraft, provides performance data from previous testing, if relevant.
7.	Special Planning	a. Special or peculiar facts established during the validation of the test requirement.
		b. Special concepts, tactics and techniques to be employed.
		c. Operating conditions relating to the requirement for test.
		d. All relevant matters not included elsewhere in the Test Directive.
8.	Participating Agencies and Responsibilities	a. Lists all agencies, organizations, and contractors who will participate in or support the test.
		b. Defines responsibilities of participants.
9.	General Project Requirements	a. Provides USAF Precedence Rating.
		b. Specifies authority to release test information and provides security guidance and direction.
10.	Milestone Schedule	Identifies milestone events and proposed date for each event.

Para. #	Heading	Content
11.	Reports	Provides guidance and direction for preparation, format, coordination and distribution of interim and final reports and as appropriate, the AFTEC/CC Summary Report.
12.	HQ USAF Project Officer	Lists the Air Staff Project Officer by name, rank, office symbol, security clearance, telephone number, and mailing address.
13.	Distribution List	Includes test participants and information addressees as appropriate. When the Test Directive is in letter form, this will be the last page.

Distribution of Test Directives. The determination of distribution requirements is the responsibility of the Air Staff Project Officer. The Test Directive will be provided to each participant (including contractors) in the test, and to each command or agency having a supervisory responsibility or monitoring interest.

Coordination. The coordination of the Test Directive is the responsibility of AFTEC and the Air Staff Project Officer. Coordination may be categorized as:

a. Coordination to validate the testing requirement.

b. Coordination to conduct preliminary planning.

c. Coordination to identify and allocate test resources.

d. Coordination in the preparation of the Test Directive, and

e. Air Staff coordination.

(1) Validating the test requirements. To the extent required, the indications that a test is needed will be studied and analyzed and previous efforts to

collect data of the type desired and the results of such efforts will be reviewed. Other factors in the validation of the requirement involve coordination between AFTEC and the Air Staff Project Office, the user command, and others as required to determine other data requirements that might be satisfied by the test; the review of available data to preclude duplication; and the review of interrelationships in the system, mission, environmental factors, enemy capabilities and tactics, and human factors relevant to the proposed test and the requirement for data.

(2) Preliminary planning. This portion of test advocacy requires the coordination of the Air Staff Project Officer, AFTEC and the user command in the preparation of a preliminary plan (based on the validation of the test requirement and the overall test purpose) which identifies critical questions and issues and defines test objectives, prescribes the necessary test environment and test milestones, and establishes data requirements and the priority of the test in relation to other programs.

(3) Identification and Allocation of Resources. This element of coordination conducted by AFTEC involves the determination of where, when, and by whom the test can be conducted and the known limitations in facilities and resources.

(4) Preparation of the Test Directive. The preparation of the draft directive by AFTEC requires the integration of the validated test requirement, preliminary planning, and resource considerations into the draft formal Test Directive and the coordination of the draft directive with the Air Staff Project Office, participating commands, contractors, and agencies to ensure that responsibilities assigned for conducting and supporting the test are understood by all parties, are concurred in, and are feasible.

For Joint Tests, this latter coordination is especially important, and shall result in a formal letter or Memorandum of Agreement between the Air Force and the participating Services/agencies. This formal agreement is appropriate for inclusion in the Test Directive.

(5) Air Staff Coordination. Prior to issue, Test Directives must be approved by the Deputy Director for OT&E. Letters/Memorandum of Agreement and Joint Test Directives will normally be approved by the Assistant for OT&E.

Test Directive Inputs and Outputs. Influences which effect the Test Directive are:

a. Input Influences:

(1) The identification of critical question/issues, the requirement for responsive data, and the significance and urgency of the resulting management decision.

(2) The availability of test resources and technology required to produce the desired data.

b. The influences resulting from the Test Directive are:

(1) The preparation of implementing directives (Project Order and Project Plan) within the major command assigned responsibility for the conduct of OT&E.

(2) The initiation of actions by all participating commands, contractors and agencies resulting in the planning, conducting and reporting of the prescribed Operational Test and Evaluation with objectives as stated in the Test Directive.

References.

a. HQ USAF Deputy Director Operating Instruction (DDOI) No. 10-2; subject: "Test Directives", dated 5 April 1973.

b. Annex D, USAF Program Action Directive 74-1; Operations, dated 15 January 1974.

Addendum. For command-initiated OT&E projects, a HQ USAF Test Directive is not issued. In such cases, the Project Order, prepared and issued by the HQ of the major command initiating the OT&E project, is the authoritative basic document for the planning, executing, supporting and reporting of OT&E. While not universally prescribed, the contents of the Project Order will generally conform to the Test Directive contents described herein, as will the coordination effort, influences, and other matters modified as required to the specific requirements of the major command conducting the OT&E project.

For categories of OT&E other than command-initiated OT&E projects (such as AFTEC conducted OT&E projects, AFTEC conducted Joint OT&E Tests and command conducted OT&E projects), a HQ USAF Test Directive is issued which is the authoritative basic document for planning, executing, supporting and reporting OT&E. In such cases, the Project Order is a document which, at the major command level, implements the HQ USAF Test Directive.

Project Orders are distributed as prescribed by the major command to include distribution (for information only) to AFTEC.

Appendix B

OFFICE OF THE SECRETARY OF DEFENSE OT&E INTERESTS

1. PURPOSE

A recurring theme throughout this manual is that OT&E is not an end in itself but rather is a systematic means of providing needed answers for the achievement of higher purposes. Fundamentally, USAF OT&E assists in providing an improved USAF mission performance posture which in turn supports the overall mission of the Department of Defense. Because of these facts, the role of the Office of the Secretary of Defense in OT&E is important to understand. The purpose of this Appendix is to summarize that role.

2. AUTHORITY

Weapon System acquisition and improvements to the existing inventory are normally thought of in the context of Research and Development Test and Evaluation (RDT&E). In this regard, the authorities and responsibilities vested in the Secretary of Defense by the National Security Act of 1947 (as amended) are unique:

> "... to eliminate unnecessary duplication in the Department of Defense, and particularly in the field of research and engineering -- its overall direction and control (is vested) in the Secretary of Defense".

A later strengthening of the intent of Congress was the passage of the Defense Reorganization Act of 1958 in which the Office of the Director of Defense Research and Engineering was established as a key element of the Office of the Secretary of Defense.

3. ORGANIZATION

a. DDR&E - The Director of Defense Research and Engineering is the principal adviser and staff assistant to the Secretary of Defense in the fields of scientific and technical matters; basic and applied research; research, development, test, and evaluation of weapons, weapon systems, and defense materiel; and design and engineering for suitability, producibility, reliability,

maintainability; and environmental services. He supervises all research and engineering activities in the Department of Defense.

(1) WSEG - An important organization that supports the DDR&E in the particular area of the concern of this manual is the Weapons System Evaluation Group. The WSEG provides the Department of Defense with operational analyses and evaluations. The Group functions under the administrative direction of the Director of Defense Research and Engineering and performs studies for the Joint Chiefs of Staff, the Director of Defense Research and Engineering, and other elements of the Office of the Secretary of Defense.

(2) DD(T&E) - To get us closer to the OT&E interests of the OSD we must introduce a key Deputy of the DDR&E. He is the Deputy Director for Test and Evaluation. The DD(T&E) has the following responsibilities:

(a) Across-the-board responsibilities for test and evaluation in the Department of Defense. Review test and evaluation policies and procedures of the DOD and the Military Departments and recommend changes as appropriate.

(b) Monitor closely test and evaluation programs conducted by the Services for DSARC programs and such other programs as he believes necessary throughout the entire testing cycle. Report to the DSARC and directly to DepSecDef his assessment as to the adequacy of the list of critical issues and problems to be attacked by test and evaluation and the schedule of test milestones, and report at Milestone III (the time of decision regarding full scale procurement of a system) to the DSARC and to the DepSecDef his independent recommendation.

(c) Insuring timely OT&E is accomplished with operational personnel in as realistic an operating environment as possible and, where practical, using pilot or early production items.

(d) Requesting Service test plans and test results as may be required to accomplish the above as early as such plans are developed by the Services and needed by DT&E.

(e) Initiating and coordinate appropriate joint testing.

(f) Overseeing the evaluation of foreign systems for possible DOD use.

(g) Administering for OSD its responsibility for the national and major Service ranges.

Other elements of the OSD that are heavily influential in the weapon system acquisition cycle (and thus directly and indirectly associated with Test and Evaluation) include:

b. Assistant Secretary of Defense (Comptroller) - The Assistant Secretary of Defense (Comptroller) advises and assists the Secretary of Defense in the performance of the Secretary's programming, budgetary, and fiscal functions and organizational and administrative matters pertaining to these functions; provides for the design and installation of resource management systems throughout DOD; and collects, analyzes and reports resource management information for the Secretary of Defense and, as required, for the Office of Management and Budget, the Congress, the General Accounting Office, and other agencies outside of DOD. He supervises, directs, and reviews the preparation and execution of the DOD budget and administers services pertaining to automatic data processing and central data services.

c. Assistant Secretary of Defense (Installations and Logistics) - The Assistant Secretary of Defense (Installations and Logistics) is the principal staff assistant to the Secretary of Defense in the fields of material requirements; production planning and scheduling; acquisition, inventory management, storage, maintenance, distribution, movement, and disposal of materiel, supplies, tools, and equipment; small business matters; transportation, petroleum, and other logistical services; supply cataloging, standardization, and quality control; commercial and industrial activities and facilities; military construction, including Reserve Forces facilities; family housing; real estate and real property, including general purpose space; and industrial relations. He is also responsible for assessing the vulnerability of resources to attack damage and for international civil emergency planning.

d. <u>Assistant Secretary of Defense (Program Analysis and Evaluation)</u> - The Assistant Secretary of Defense (Program Analysis and Evaluation) reviews, for the Secretary of Defense, quantitative requirements including forces, weapon systems, equipment, personnel, and nuclear weapons; assists the Secretary in the initiation, monitoring, guiding, and reviewing of requirements studies and cost-effectiveness studies; encourages the use of the best analytical methods throughout the DOD; and conducts or participates in special studies as directed by the Secretary of Defense.

4. KEY OSD ASPECTS IN SYSTEM ACQUISITION

Without belaboring the details of the current process of acquisition of major weapon systems at least two aspects of this process need be cited here as being key.

The first is a document known as a Development Concept Paper which is basic to the process.

The DCP is prepared to pull together and establish agreement on the issues and considerations which should go into SECDEF's go no-go milestone decision. Once approved by SECDEF, the DCP constitutes a "contract" between the Secretary of Defense and the Service for the decentralized management of the program, and also serves as a concise statement of program rationale.

The DCP, including all appendices, is held to a concise 20-page limit to keep it to a length that can be read and assimilated by busy high-level officials, including the Secretary of Defense.

Each DCP is tailored to both the program and the decision for which it is prepared, thus each is to some extent unique. However, most cover the following subjects:

- The decision the Secretary is being asked to make
- The alternatives which should be considered in reaching this decision
- The reasons for having the program, i.e., the threat or other condition which constitutes the "demand" for the program.
- Business/Management plan for carrying out the program

- Projected costs
- Risks and measures for dealing with them
- The limits of the grant of authority from SECDEF to the Service to manage the program. Thresholds are normally set on costs, schedule, particularly for program milestones, and demonstrations of performance.
- Test and evaluation

Before introducing the second key aspect it will be constructive at this point to digress and discuss Test and Evaluation to assure a common understanding. This point is chosen for this discussion because immediately above it is noted that the all important DCP requires a section on T&E.

While the terms "test" and "evaluation" are most often found together, they actually denote clearly distinguishable functions. "Test" denotes the actual testing of hardware/software — models, prototypes, production equipment, computer programs — to obtain data of value in developing new capabilities, managing the process, or in making decisions on the allocation of resources. "Evaluation" denotes the process whereby data is logically assembled and analyzed to aid in making systematic decisions.

T&E is the deliberate and rational generation of data concerning the nature of the emerging system and the creation of information useful to the technical and managerial personnel controlling development. In the broad sense, T&E may be defined as all physical testing and experimentation and related analyses performed during the course of the research, development, introduction and employment of a weapon system or subsystem. T&E is continuous throughout the system's entire life although it has various labels on it at different times and different purposes under these labels. For example there is Development T&E and Operational T&E which in turn is divided into Initial OT&E and Follow-on OT&E, Joint T&E as well as others. Again, these differences are covered in detail elsewhere and will not be treated here.

The second key aspect to the acquisition process is a formally established group known as the Defense Systems Acquisition Review Council. The DSARC membership includes those SECDEF staff officers cited above (the DDR&E, the ASD (Comptroller), the ASD (Installations and Logistics) and the ASD (Program Analysis and Evaluation). The Deputy SECDEF is a frequent participant in DSARC sessions.

5. BACKGROUND AND CURRENT INFORMATION

Some background on the OSD current high interest in Test and Evaluation including Operational T&E is valuable.

The weapon system acquisition procedures that evolved in the decade of the Sixties had several problems. Funds for testing were often neither adequate not properly "fenced" to protect them from diversion to other uses. Test results were thus seldom of good or sufficient quantity or quality or timely enough to really influence decisions on major program. Thus, costly mistakes and debilitating controversies often occurred.

Efforts started in the late Sixties to correct these problems were accelerated by the Report of the Blue Ribbon Defense Panel of July 1970. Although the recommendations of the BRDP were not implemented to the letter they were, in fact, implemented in their intent.

In his memo of 11 February 1971, "Conduct of Operational Test and Evaluation," Deputy Secretary of Defense Packard wrote: (emphasis added)

> "Although each Service now has a somewhat different way of organizing for operational test and evaluation, it is apparent to me that this function can best be performed by an agency which is <u>separate and distinct from the developing command</u> and which <u>reports</u> the results of its test and evaluation efforts <u>directly to the Chief of the Service.</u> Moreover, within the Service headquarters staff, there needs to be an office with a clear OT&E identification to provide staff assistance directly to the Service Chief and to provide a headquarters focal point for the <u>independent OT&E field agency</u>. Thus, at the completion of Operational Service Testing in the Army, OPEVAL in the Navy, and Cat III testing in the Air Force, I would expect that the respective Service Chiefs

would have a clear picture of the operational suitability of a weapon system for Service use, to include its principal deficiencies and limitations and the corrective actions required prior to full-scale introduction into the force. Accordingly, each Service is requested to restructure its organization for OT&E along the lines specified above."

"As a second step, I am establishing a Deputy Director for Test and Evaluation within ODDR&E with across the board responsibilities for OSD in test and evaluation matters. This office will review and approve test and evaluation plans prepared by the Services and will provide an assessment of results obtained."

Prior to the establishment of the ODD(T&E) the ODDR&E was not completely without some OT&E influence. An Assistant Director for OT&E was established in 1966. This office became part of the ODD(T&E). Further, an Assistant Director for Ranges and Space Ground Support was established in 1960. This office had the ODDR&E - staff level and responsibility for the major DoD Test Ranges. Similarly, the OAD(RSGS) (with a new function of providing targets) became a part of the ODD(T&E). The evolution resulted in the current structure of the ODD(T&E):

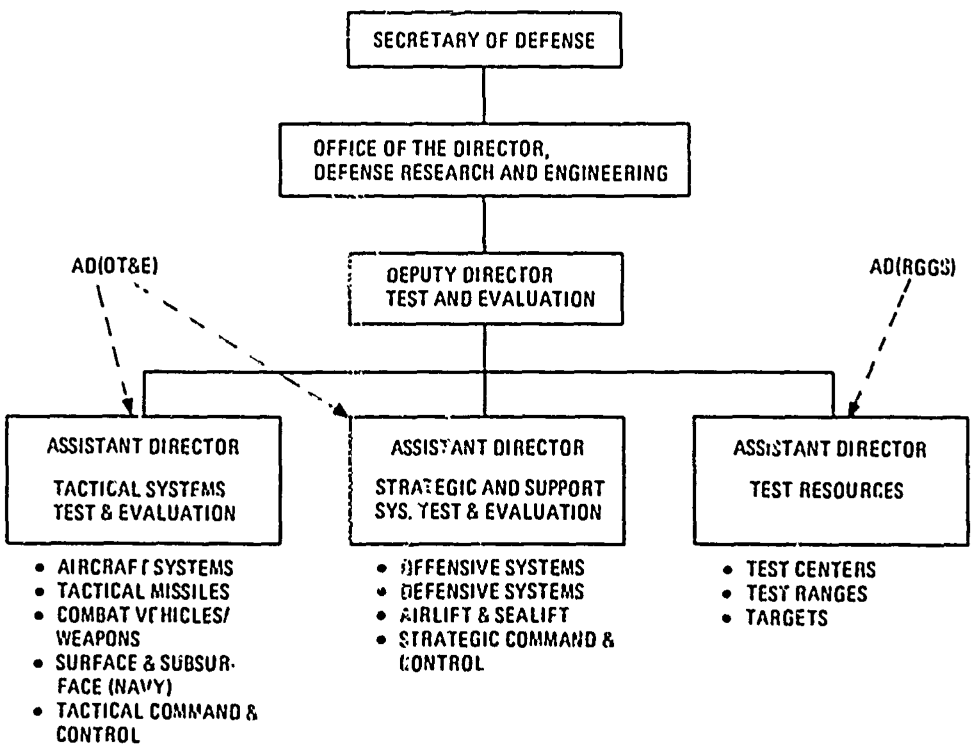


Figure B-1. Structure for OT&E within OSD

As the Defense Department worked to improve its OT&E posture, the Congress also took an important step.

The Armed Forces Appropriation Authorization Act for FY 1972 (Public Law 92-156) included a Section 506 which stated (in part):

"(a) Beginning with 1972, the Secretary of Defense shall submit to the Congress each calendar year, at the same time the President submits the Budget to the Congress, a written report regarding development and procurement schedules for each weapon system for which fund authorization is required and for which any funds for procurement are requested in such budget. Beginning with the calendar year 1973, there shall be included in the report data on operational testing and evaluation for each such weapon system for which funds for procurement are requested.

(b) Any report required to be submitted under this section, shall include detailed and summarized information with respect to each weapon system covered by such report, and shall specifically include, but shall not be limited to--

(1) To the extent required by the second sentence of subsection (a) of this section, the results of all operational testing and evaluation up to the time of the submission of the report, or, if operational testing and evaluation has not been conducted, a statement of the reasons therefor and the results of such other testing and evaluation as has been conducted."

The important thing to note regarding Section 506 is that if a good story cannot be told about OT&E on any given system no Procurement Funds will be authorized and appropriated. This excellent and not exactly subtle incentive caused many to become very interested in OT&E.

The steps taken by the Military Departments to improve their OT&E posture are detailed in subsequent sections of this manual.

Several powerful and effective tools have been provided to the DD(T&E) that enable him to carry out his responsibilities. DoD Directive 5000.3 "Test and Evaluation" (January 1973) is really the possession of the DD(T&E). It is his charter and authority to act. It states the policy framework within which he functions and upon which all Military Department regulatory documents on T&E (e.g., AFR 80-14) are based. It is thus the cornerstore of all T&E carried out in the entire Department of Defense (including its component Departments of Army, Navy and Air Force). It is "must" reading for all who are associated with DoD T&E.

A further tool is a separate and distinct appropriation of funds directly to him for his control and apportionment to the various programs and uses that he cited in his request to the Congress for the funds. These funds are contained in Program Element 65804D. They are from the overall Research Development Test and Evaluation Appropriation and thus have the flexibilities associated with the RDT&E Program. The annual sum has historically been between $25-30 million.

Joint Testing - Recall the earlier listing of the responsibilities of the DD(T&E)--specifically he is to "initiate and coordinate appropriate joint testing". Now, when the DD(T&E) recognizes a need and after due study, discussion, and coordination, he causes a joint test to be directed on two or all of the Military Departments (as participants) somebody has to have the money to pay for the test costs. It is somewhat impractical, to say nothing of being a source of irritation, to designate one of the Departments to "cough up the money". Thus, 65804D funds are, for the most part, designated to pay these costs--at least the "unique" or "test peculiar" costs. The participating departments do absorb the costs of their manpower and other not easily specifically identifiable costs (O&M for participating aircraft for example). By providing funds it is also easier for the DD(T&E) to maintain control of the tests.

Several such joint tests have been and are currently being conducted. Examples include:

a. The Joint Electronic Warfare Test: With the Navy as the designated Executive Agent, this test is to provide information on electronic warfare tactics, techniques and equipment effectiveness under realistic operational conditions.

b. The Hit Probability Joint Test: Here the Army is executive. The tests are to provide an empirical basis for predicting U.S. aircraft attrition in future tactical engagements. The data obtained will modify and give greater confidence to attrition prediction models which have never been validated by actual measured data.

Others, simply by title are:

c. Laser Weapons Joint Test

d. Laser Guided Weapon Countermeasures Joint Test

e. Airborne Target Acquisition Joint Test

f. Close Air Support Command & Control Joint Test

g. Radar Bombing Accuracy Joint Test

h. Air to Air Weapons Joint Test

i. Aircraft Survivability Joint Test

Although the bulk of the funds appropriated to the DD (T&E) Appropriation is used to support joint tests other uses can be cited. For example certain T&E related studies have been done, a "level-of-effort" consultant service is obtained from one or more FCRC houses, certain "get-well" money was provided to each Military Department early in this OT&E resurgence era, and funds for studies associated with the definition of the Continental Operations Range were provided to the Air Force.

Other Involved OSD Elements - OSD interests in OT&E, as noted earlier, are also held by offices other than the ODDR&E/ODD(T&E). For example the basic policy memorandum (15 Nov 72 Memorandum for Assistant Secretaries of the Military Department's/(Financial Management) from the Assistant Secretary of Defense, Subject: "Congressional Data Sheets Required by Section 506, PL 92-156, in support of the FY 1974 Authorization and Budget Requests") that implemented Section 506 of PL 92-156 was issued by the Assistant Secretary of Defense, Comptroller. This memo expanded the words of Section 506 to the point where it is possible to prepare

a "check list" to assure compliance with the Congressional requirement for OT&E information. The topics to be addressed are:

a. Narrative on OT&E accomplished to date.

b. Additional OT&E to be accomplished prior to program budget year major production contract award.

c. Where applicable, why the IOT&E required by current DOD Policy will not have been completed by that date.

d. Summary description.

e. Schedule.

f. IOT&E/FOT&E results to date.

g. Degree of similarity between item tested and item to be procured.

h. Major subsystems not included in the testing.

i. Agency responsible for OT&E.

j. State whether an independent or combined test.

k. Location where conducted.

l. Type of personnel who operated and maintained the item during test.

m. Major discrepancies.

n. Remedial action to be/or taken.

o. Planned retesting/schedule.

p. Maintainability/reliability comments.

q. Identify IOT&E not accomplished prior to production contract decision, with explanation.

r. Germane operational experience applicable to item operational suitability/ effectiveness.

Still another step taken under the aegis of the ASD(C) was the issuance of a new funding policy for the various elements of the DoD Test and Evaluation Facility Base. (The TEFB and its 26 elements is discussed elsewhere in this manual). All 26 elements do not come under the new policy. Those omitted were considered to provide services only to their parent Military Department and thus the choice of a method of funding was left to the applicable department.

The policy basically requires a user of the range, test facility or test center to pay for the direct costs ascribed to its use. All other costs of maintaining, operating, developing and retaining a modern, viable capability for test support is provided to the range, test facility or test center directly as an institutional fund.

This policy is specifically mentioned to warn the Test Director reader of this manual to specifically investigate the funding procedures of any such place he may plan to use well in advance of his planned use. If this is neglected it is easily conceivable that an unplanned expense could manifest itself in a highly embarrassing fashion when a "range use" bill is presented.

Recall the DSARC discussion above and the members of that body. On each major program the DSARC must meet at least three times. These meetings occur between each of the prescribed phases of the Acquisition Cycle as follows:

Conceptual Phase	
	DSARC I
Validation Phase	
	DSARC II
Full Scale Development Phase	
	DSARC III
Production & Deployment Phase	

Since meetings I and II are largely concerned with development of the system they are more the province of the DDR&E. Thus, he chairs those sessions. Meeting III however, is where the big step is taken that commits the DoD to the spending of large amounts of money for production of N-units of the system. In fact, to even qualify for DSARC attention at all the predicted production costs must be at least $200 million. DSARC III is chaired by the Assistant Secretary of Defense (Installations and Logistics).

The ASD(I&L) is, of course, very concerned about the OT&E performed in connection with all such major acquisitions. It is at DSARC III where the OT&E stops being "Initial" and becomes "Follow-on". Thus the ASD(I&L) is the key figure at this point in the acquisition cycle.

A reference back to the responsibilities of the ASD(I&L) cited earlier shows he is concerned about such matters as system maintenance, supplies, tools, and other logistic functions. All of these must be tested (and evaluated) "with operational personnel in as realistic an operating environment as possible"--and this is indeed OT&E.

Finally, we must talk about the OT&E interests of one other OSD official not previously mentioned. This is the Director for Telecommunications and Command & Control Systems, This Director (who reports directly to the SECDEF) was, in the past, titled Assistant Secretary.

Note the previously given organization of the ODD(T&E). The areas of concern of each of the Assistant Directors are listed under each AD box. Communication systems are not mentioned.

The Director, TCCS has total cognizance over the systems (primarily communications) he is responsible for including development and test and evaluation. There is, of course, significant coordination and interface between the Director and the DD(T&E). Fundamentally however, the Director is the responsible party when it comes to OT&E of such systems.

6. SUMMARY

In summary the Secretary of Defense, his Deputy, his Director of Research and Engineering and his Assistants for Comptroller (functions), Installations and Logistics, Program Analysis & Evaluation and his Director for Telecommunications and Command & Control Systems all are highly concerned about OT&E and thus exert major influence on its conduct. The key man is the DD(T&E). The policies and procedures regarding OT&E coming from the office of the SECDEF must be read, understood and complied with. In so doing the fundamental and higher purpose of OT&E of assisting in an improved defense posture is served.

Appendix C

ARMY OPERATIONAL TEST AND EVALUATION - AN OVERVIEW

1. INDEPENDENT TEST AGENCY FOR OT&E

OTEA. The U.S. Army Operational Test and Evaluation Agency (OTEA), a Field Operating Agency (FOA) reporting directly to the Chief of Staff of the Army, is the Army's independent test agency for operational test and evaluation (OT&E). OTEA is located on North Post, Fort Belvoir, Virginia.

2. BACKGROUND

Pre-OTEA Testing. During the period 1962-1971, the testing of systems and equipment under development and in-process of acquisition by the Army was accomplished almost totally within the Army Material Command (AMC). Within AMC the "material developer," a Project Manager or equivalent conducted Engineering Development Tests (EDT), R&D Acceptance Tests (RDAT) and similar development tests. The focal point for test activity within AMC was (and for DT&E continues to be) Test Command (TECOM), which is independent of the material developer, but under the command of AMC. TECOM employed both Engineering Tests (ET) and Service Tests (ST) in the testing and evaluation of systems and equipments. ETs were conducted by TECOM to determine if systems met the material need. STs examined the human interface with the system, were operational testing oriented, and were conducted within AMC under the appropriate Board (artillery, infantry, aviation, armor, engineer, air defense, airborne and communications) at a test center. These tests employed typical user troops in a tactical environment and generally possessed many of the characteristics of present day OT&E. Service tests conducted in this era had certain disadvantages, however. These tests were often conducted after the management decision to produce the emerging system/equipment in numbers, and even after the deployment for Army operational use. Such testing was funded by the developer, was subject to cost-cutting, and was not independent of the influence of the developer.

Commencing with the Deputy Secretary of Defense (DEP SEC DEF) memorandum to the military components of the DOD in May 1970, the Department of Defense made several sweeping policy changes concerning the acquisition of material and the role of test and evaluation in providing timely information to management on which to base development and acquisition decisions. Ultimately, these policy changes were formally set forth in DOD Directives (DODD) 5000.1, "Acquision of Major Systems" and DODD 5000.3, "Test and Evaluation". These directives impacted the systems acquisition procedures, test and evaluation practices, and organizations in each military service, and within the Department of the Army, resulted in significant changes in policy and organization.

The initial response of the Army to the DOD Directives was to attempt to satisfy the new material acquisition and test policies by reorienting the Army acquisition policies and procedures to conform to the new DOD directives and relocating some OT&E responsibilities such as the "Intensified Confirmatory Troop Test" (ICTT) from TECOM to Combat Developments Command (CDC). Other OT&E, such as the "Expanded Service Test" (EST) remained with TECOM. This location of OT&E functions within an existing command in the Army structure proved to be only an interim measure. Before CDC's assigned responsibilities in OT&E could be fully implemented, a reorganization of the Army was effected, including the decision to establish a truly independent field agency responsible for OT&E. This decision is reflected in a Department of the Army (DA) Memorandum dated 20 June 1972; subject: "Material Acquisition Guidelines" which was implemented by a DA Letter of Instructions (LOI) dated 23 August 1972.

3. FORMATION OF OTEA

On 22 September, 1972, an Army Adjutant General Letter (AGGL) was promulgated, establishing OTEA with a level of manning at 53 officers, 2 enlisted men, and 65 civilian personnel. Over the following four months, the OT&E responsibilities in CDC for major and certain non-major systems were transferred to OTEA, with the remaining responsibilities in operational testing of nonmajor systems relocated in the newly organized Training and Doctrine Command (TRADOC).

On 15 January, 1973 OTEA reported having attained the full operational capability to prosecute mission, tasks, and functions assigned.

The significant dates in the history of the Operational Test and Evaluation Agency are:

23 August, 1972 - Issuance of Department of the Army LOI which established the policies on which OTEA functions are based.

22 September, 1972 - Promulgation of Army Adjutant General Letter which was formally established OTEA.

15 January, 1973 - OTEA became fully operational.

Mission and Tasks. The mission and functions of OTEA are assigned by Army Regulation (AR) 10-4, "U. S. Army Operational Test and Evaluation Agency" and are as follows:

a. Mission. "The mission of OTEA is to manage all user testing, operational testing (OT), force development testing and experimentation (FDTE), and joint user testing directed by the Office of the Secretary of Defense (OSD)."

b. Specific Tasks. Specific tasks are assigned to OTEA which are summarized as:

(1) Plan, direct, and evaluate the operational testing of all major and selected nonmajor systems.

(2) Coordinate the operational testing of other nonmajor systems.

(3) Program and budget the requirements financed under Operations and Maintenance, Army (OMA), Program 2, and coordinate funding for requirements to be financed by all other appropriations for assigned testing.

(4) Manage major and coordinate nonmajor FDTE.

(5) Coordinate Army participation in the planning and execution of Joint OT&E.

(6) Provide a strong focal point organization at HQDA to keep DA and OSD fully informed on the Army's OT needs and accomplishment.

c. Assigned Functions. OTEA is assigned specific functions in the area of user testing. User testing includes Operational Testing (OT), Force Development Testing and Experimentation (FDTE), and Joint Testing. (See paragraph 4 for an explanation of these terms.) The functions in user testing set forth below are assigned to OTEA:

(1) Operational Testing (OT)

(a) (All major and selected nonmajor systems in acquisition) Plan, program, budget (oridentifybudget requirements, as appropriate), schedule OT, prepare an independent evaluation which assesses the significant operational issues and the adequacy of testing for all such systems.

(b) (Other nonmajor systems) Manage testing and assist the DA proponent, when required, in assessing the quality of test plans, execution, and reports.

(2) Force Development Testing and Experimentation (FDTE)

(a) (All major FDTE) Manage testing, to include: programming; budgeting; scheduling; approving test design of selected test plans; monitoring the conduct and ensuring the quality of testing; providing a review and evaluation of test results to HQDA and others as required. Coordinate the planning, programming, budgeting and scheduling with the Test Schedule and Review Committee (TSARC).

(b) (Other nonmajor FDTE) Manage testing and assist the DA staff proponent as required to assure the quality of test plans, test execution and reports.

(3) Joint Testing. Assume test management responsibilities for OSD sponsored tests involving the Army, to include:

(a) Initial point of contact for the Army.

(b) Program, budget (or identify budget requirements, as appropriate).

(c) Schedule and coordinate Army participation.

(d) Recommend to HQDA the Army command or agency to function as the proponent for doctrinal and technical aspects of Army participation.

(e) Assist the Army Test Director by providing testing expertise and budget planning.

(f) Review, comment on, recommend for approval the overall test plan

(4) Other Functions Assigned to OTEA. Focal point for DA for the management of user tester instrumentation and coordinator of financial planning and management of all OT, FDTE, and joint and combined testing, to include:

(a) Preparation and submission of user testing documentation (including justification) to support development of the Army Program Objectives Memorandum (POM).

(b) Developing recommended funding levels and mission priorities for DA.

(c) Reviewing, validating, prioritizing and justifying all user test requirements submitted to DA during the Budget Development Cycle.

(d) Monitoring and reviewing the utilization of user testing funds and recommend revisions, as appropriate, based on changes to missions and priorities.

Command Relationships. The Commanding General OTEA reports to the Chief of Staff of the Army and has an alternate office in the CSA office to provide the

headquarters focal point required in response to the Deputy Secretary of Defense memo of 11 February 1971. Other OTEA relationships are prescribed as follows:

a. OTEA is authorized direct communications and liaison within HQDA and elsewhere as required in the prosecution of assigned user testing responsibilities.

b. OTEA functions as the Army point of contact for user test matters with ODDRE (T&E) and OSA.

c. In the tasking of Army commands and agencies in matters concerning user testing, OTEA prepares the appropriate assignment documents for the CSA and submits such documentation through the normal chain of command.

4. OTEA INVOLVEMENT IN THE ACQUISITION CYCLE

Preface. There are significant differences between the military services in the organization and procedures employed by each in their response to the DOD Directives concerning material acquisition and testing. For the Air Force reader, these differences must be identified and explained in order that valid comparisons may be made for an appreciation of the practices within each Service.

In order to understand OTEA's role, certain characteristics of Army material acquisition and tests are provided as follows:

Army Systems/Equipment Testing. The testing of Army systems or equipment is divided into Development Testing (DT) and User Testing, the elements and characteristics of which are:

a. Development Tests (DT). DT is essentially the same throughout all the military services. These tests are technical requirements-oriented, managed by the material developer and conducted throughout the development cycle to determine primarily the degree to which an item meets performance specifications. Army DT is usually conducted in three phases as:

(1) DT I - Conducted in the validation phase to assess technical risks and support the request for program transition to the full-scale development phase.

(2) DT II - Conducted in the full-scale development phase to resolve or minimize design risks problems, provide an estimation of the system/ equipment military utility, and support the request for program transition to limited production (occurring in the latter full-scale development phase) or when appropriate, full-scale production (the production phase).

(3) DT III - Conducted to ensure that deficiencies observed in DT II are corrected in production equipment.

b. User Testing. User testing is a generic term encompassing Operational Testing (OT), Force Development Testing and Experimentation (FDTE), and Joint Testing.

(1) Operational testing (OT) - OT is the testing of systems in the material acquisition process by an organization independent of the developer. As with DT, operational testing is usually conducted in three phases as:

(a) OT I - Conducted in the validation phase usually in combination with DT I, to provide an indication of utility and worth to the user, identify critical operational issues, and provide information to support program transition into full-scale development.

(b) OT II - Utilizes pre-production equipment in troop unit field exercises to examine the resolution of critical operational issues. Conducted in the full-scale development phase, occasionally in combination with DT II.

(c) OT III - Examines initial production items to ensure that operationally critical issues are resolved and that the item/system is operationally suitable. Conducted in the full-scale development phase, separate from DT III.

(Figure C-1 depicts the Army DT and OT testing in the material acquisition cycle. This illustration shows the usual combination of DT I and OT I, the normally separate conduct of DT II and OT II, and DT III, and OT III, and the phase in which tests are conducted.)

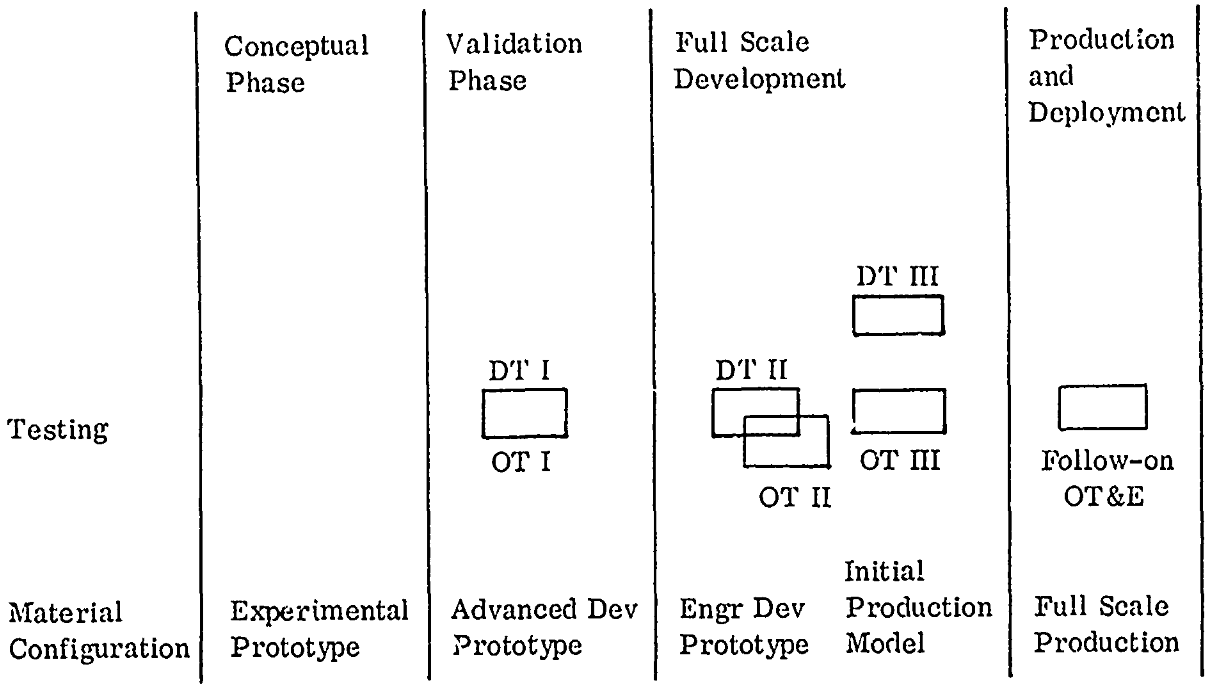


Figure C-1. Testing in the Army materiel acquisition cycle

(DT II, OT II, and DT III is conducted on all development items/systems as a minimum. Non-developmental items normally undergo only DT III.)

(2) FDTE - This element of user testing consists of tests ranging from field experiments (small scope, highly instrumented, high resolution) to field tests (broad scope, less instrumented, low resolution, highly subjective). FDTE assesses interdependence among doctrine, tactics, organization and material and often supports the acquisition process. Its purpose is also to develop or demonstrate the adequacy of doctrinal, organization training and logistical concepts. FDTE is categorized as "major" or "minor" and is further subdivided into:

(a) Field experiments

(b) Field tests

(c) User evaluations

(d) Military potential tests

(3) Joint Tests - An element of user testing conducted to evaluate Army systems or concepts having an interface with or requiring a test environment of another Service, or systems or concepts of another Service which require testing in an Army environment.

5. THE REQUIREMENTS PROCESS

The degree of OTEA's involvement in the acquisition cycle depends upon whether the system in acquisition is designated a "major system," "selected nonmajor system" or "nonmajor system." (Definitions for these terms are provided in this document.) Within the Department of the Army (DA), as with the Air Force, the Required Operational Capability (ROC) establishes the need for an operational capability. The submission of the ROC is the first step in the development of a system. The ROC is submitted to the Deputy Chief of Staff for Operations (DCSOPS) by the originator. DCSOPS is responsible for ensuring that the ROC is adequate and for coordinating the ROC among the Army Staff.

DCSOPS, in coordination with the Deputy Chief of Staff Research Development and Acquisition (DCSRDA), will determine whether the system proposed by the ROC is to be considered a major system or a nonmajor system. Major systems include not only those which qualify for a Defense System Acquisition Review Council, but also other systems, designated as "selected nonmajor systems," which are critically important to the Army, are complicated, expensive, controversial, or which for other reasons require the involvement of the top management level of the Army.

Major Systems. For major or selected nonmajor systems, DCSOPS is responsible for processing the ROC for approval, and following approval, setting in motion the special management policies and procedures reserved for such systems. This special management effort includes the establishment of a Special Task Force in the conceptual phase which is responsible for producing the draft Development Concept Paper (DCP), the Concept Formulation Package (which supports the content of the DCP), and a Final Report which is provided to the material developer (the Project Manager or equivalent office). The Final Report forms the basis for and contains

necessary guidance and inputs to the Development Plan (DP), prepared by the Project Manager. The plan for test and evaluation is contained in the Final Report and when translated into the DP by the Project Manager, it becomes Section IV of the DP entitled "Coordinated Test Program."

Nonmajor Systems. The ROC for a nonmajor system is approved by DCSOPS in coordination with DCSRDA. In addition, DCSRDA designates the using command or agency with which the system developer (usually the Army Material Command (AMC)) will coordinate acquisition, including the development, testing, production and deployment and is directly responsible for the General Staff management of the development of other than major systems. Prior to authorizing the development (and depending on the program importance and risks) the DCSRDA may direct the construction of a Concept Formulation Package and a Development Plan by the material developer, with assistance as appropriate, by the user, logistician, and test commands and agencies. Nonmajor acquisition programs are subject to such reviews and controls as the DCSRDA deems necessary. For nonmajor programs, the decision milestones are called In Process Reviews (IPRs).

Decision Levels. Figure C-2 depicts the levels of decision for major programs and nonmajor programs within the Department of the Army.

6. THE ROLE OF OTEA

In the acquisition cycle of major and selected nonmajor systems, OTEA's functions are shown in Figure C-3 and are reviewed herein in five areas, listed as follows:

a. The OTEA independent evaluation plan.

b. OTEA's participation in the Special Task Force.

c. The Coordinated Test Program.

d. Planning and conduct of field tests.

e. OTEA's independent evaluation.

TYPE ACQUISITION	LEVEL OF APPROVAL FOR MAJOR DECISIONS	DECISION BODY	DECISION RECORDING DOCUMENT	GENERAL MONETARY THRESHOLDS (MILLIONS)	REMARKS
MAJOR PROGRAMS	SECDEF	DSARC & ASARC	DCP	RDTE 50(+) / 200(+) PEMA	- Special reviews may be held to resolve issues that are delegated by the ASARC/DSARC review process.
	SA/CSA	ASARC	ARMY PROGRAM MEMORANDUM	As Directed	- Special reviews may be held to resolve issues that are delegated by the ASARC review process.
NON-MAJOR PROGRAMS	HQDA DCSRDA	IPR	SECTION I, DEVELOPMENT PLAN	As Directed	
	COMMAND (Materiel Developer)	IPR	SECTION I, DEVELOPMENT PLAN	RDTE 0-50 / 0-200 PEMA	- IPR recommendations will be approved by the materiel developer. - Formal HQDA participation limited to annual budget review, unless otherwise directed.

Figure C-2. Levels of decision (Army)

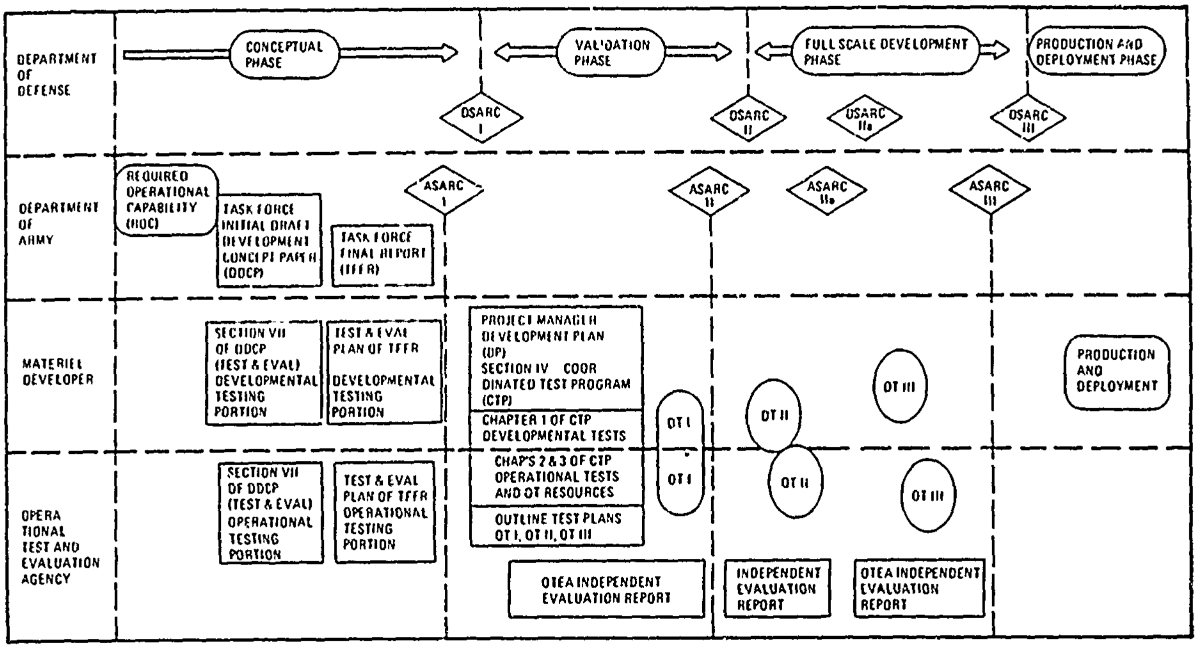


Figure C-3. OTEA Actions

The OTEA Independent Evaluation Plan. Upon the assignment of testing responsibility of a system to OTEA, usually in the conceptual phase of the acquisition cycle, the Agency begins to develop the Independent Evaluation Plan. This plan serves as a road-map for all further actions by OTEA, commencing with the OTEA participation in the Special Task Force. The objective of the plan is to document the data sources required for an evaluation of the system's operational effectiveness which is independent of evaluations of technical effectiveness, cost, and other considerations.

The plan includes all OT&E test and coordination requirements and is initially developed concurrent with OTEA's participation on the Special Task Force. The plan is updated as required during the development process.

OTEA's Participation on the Special Task Force. Upon approval of the ROC for a major or selected nonmajor system, the DA appoints a Special Task Force, an ad hoc group of experts composed of members from all major commands and agencies concerned. OTEA is included for expertise in the area of OT&E. The deliberations of the Task Force are based on analysis, tests, field experimentation and studies resulting from DA tasking. Upon determination that the development effort prescribed

by the ROC is feasible and needed, the Task Force produces three documents which have key roles in the development effort: the draft Development Concept Paper (DCP), the Concept Formulation Package (CFP) which supports the content of the DCP, and a Final Report which is provided to the material developer (the Project Manager). OTEA representatives on the Task Force provide the list of operational issues, identify operational test facility and resources requirements and provide the OT portion of the draft DCP. OTEA also provides the OT inputs to the plan for test and evaluation in the Final Report.

Normally the Special Task Force will be dissolved not later than the completion of DSARC I (the program decision milestone), at which time the Project Manager will assume management responsibility for the acquisition of the system.

The Coordinated Test Program (CTP). The Final Report produced by the Task Force and provided to the Project Manager is the basis for and contains essential elements of the Development Plan (DP). The DP prescribes the program and schedule for all further development of the system to include all DTE and OTE requirements. The DP is prepared in the validation phase by the Project Manager.

OTEA's involvement with the Project Manager (PM) commences upon selection of the PM at the start of the validation phase at which time the Agency provides the OT&E input to the Coordinated Test Program (CTP). The CTP (Section IV of the Development Plan) is the key management document for establishing the DT and OT to be accomplished by the material developer and the command/agency responsible for OT, and is the basis for the planning, coordinating, conducting, analyzing, and reporting of testing in the materiel acquisition cycle. The functions of OTEA in the CTP are:

a. Preparation of Chapter 2 (operational testing) and costing estimates.

b. The preparation of Outline Test Plans (OTP) for operational tests to be conducted. (The OT test requirements provided by the OTEA input to the Final Report forms the basis for the Outline Test Plans.

c. The incorporation of all Outline Test Plans of the Coordinated Test Program into the Army Five Year Test Plan.

d. Continual updating of the OT portion of the CTP.

Planning and Conduct of Field Tests. Commencing in the validation phase, OTEA prepares a test design plan for each operational test within the Coordinated Test Program. The test design plans are provided to the testing command prior to the start of detailed test planning at the test site. In preparing the test design plan, OTEA utilizes the OTP's and the test support data (such as proposed doctrine, organization, threat, logistics procedures and training plan) provided by supporting commands and agencies in response to requirements in the Five Year Test Plan. OTEA also effects coordination of the test design plan with appropriate commands and agencies.

Elements of test planning, other than the basic test design, are prepared by the ad hoc test directorate formed for a given test. These additional elements include test analysis planning, test evaluation procedure, data collection procedure, an outline of the test scenario as a part of a larger control plan, personnel planning (requirements and assignments, including a plan for the test directorate structure), support planning, and the test budget.

In addition to providing the test design plan, OTEA provides a field test team to fill key positions in the test directorate and to organize the test directorate at the test site. The field test team is assigned to the test directorate for the duration of planning, conduct, and reporting of the test. OTEA's role in a typical test directorate, (shown on page C-15 in Figure C-4) is to furnish 3 to 5 personnel to fill key positions such as the Deputy Director for OT, the Chief Analyst, the Chief Data Collector, and the Chief Controller. TRADOC, representing the user, provides the Deputy Test Director for Systems Concepts with appropriate assistants in areas of conceptual expertise. All other positions are filled by the testing command.

For OT conducted by Modern Army Selected Systems Test, Evaluation, and Review (MASSTER) for OTEA, there are exceptions to the foregoing: (1) MASSTER prepares the test plans, subject to OTEA approval, and (2) OTEA's participation in the field testing is usually limited to a monitor/observer team.

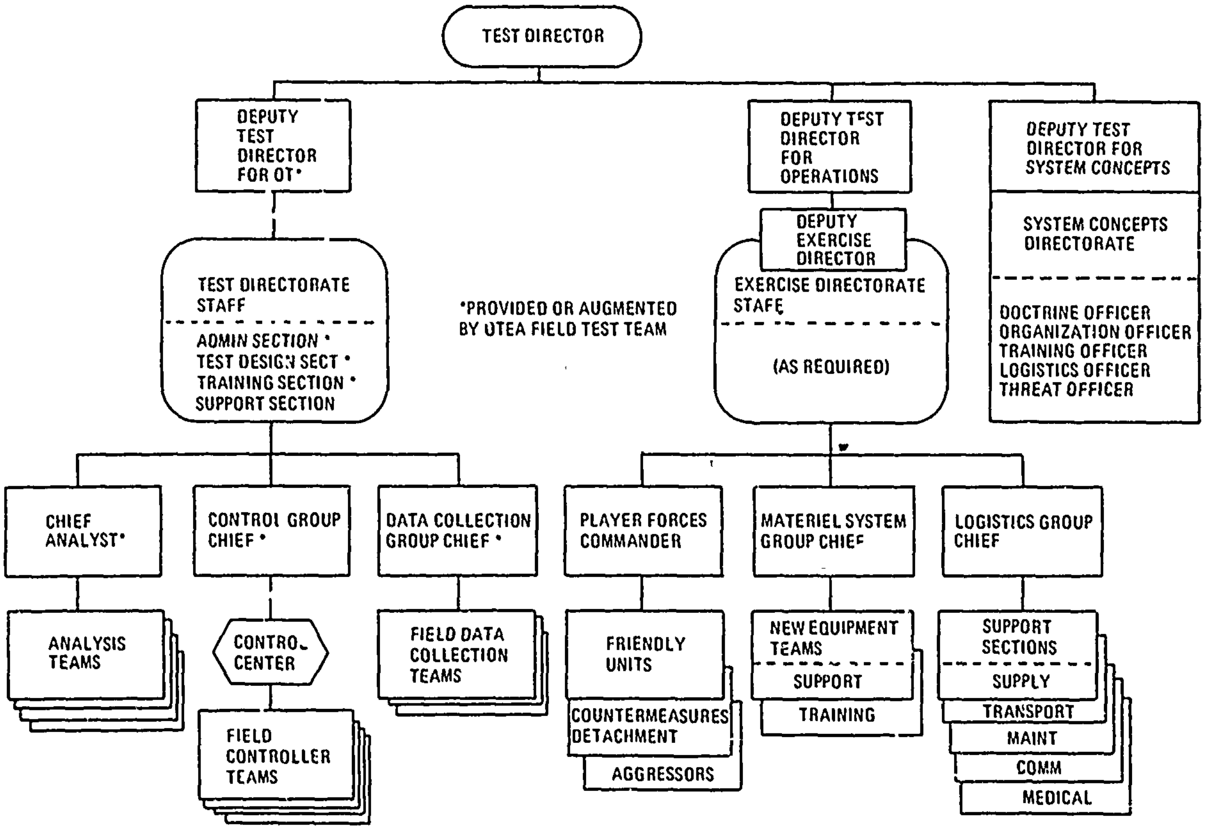


Figure C-4. Typical test directorate organization depicting OTEA involvement

OTEA's Independent Evaluation. The independent evaluation of operational effectiveness commences with the OTEA Independent Evaluation Plan in the conceptual stage of the acquisition cycle. Based on the plan, all data elements and data sources required to satisfy test objectives are identified and test objectives are continuously reviewed to ensure that objectives respond to operational issues. The evaluation utilizes all available information, including the results of development testing. OTEA's independent evaluation is provided directly to the Army Systems Acquisition Review Council (ASARC) and the Defense Systems Acquisition Review Council (DSARC), as appropriate, to assist in the decisions concerning the commitment of resources for full-scale development and limited or full production (ASARC/DSARC II, IIa, III respectively). OTEA evaluations of nonmajor systems, which involve OT

planned and conducted by other commands in coordination with the material developer, is accomplished on a case-by-case basis as directed by the Department of the Army.

7. COMBINED DT&E/IOT&E

During the conceptual phase of a major or selected nonmajor system, the documents produced by the Special Task Force tentatively determine whether DT&E and IOT&E (consisting of DT/OT I, II and III) will be conducted as combined tests. This matter is addressed in summary fashion in the draft Development Concept Paper and in greater detail in the plan for test and evaluation in the Final Report of the Task Force. The combined test form is further defined in the Coordinated Test Program of the Development Plan and gains Army authorization for implementation upon approval of the Outline Test Plans for testing of material in acquisition and the inclusion of Outline Test Plans in the official, authoritative document for all Army test programs, the Army Five Year Test Plan. The influences which drive the decision to combine DT&E and IOT&E are:

a. The development and operational issues involved.

b. The objectives of both DT&E and IOTE.

c. The availability of prototype equipment.

d. The cost in terms of resources, including time.

As a general rule, these influences dictate a normally combined DT I/OT I; the occasional combination of DT II/OT II, and rarely, the combining of DT III/OT III.

Objectives and Roles in Combined DT/OT. The following DT/OT tasks depict the objectives and the roles of both the developing agency and OTEA in the combined DT&E/IOT&E projects noted above:

a. Task: Combined DTI/OTI (validation phase)

Objectives:

DTI: (1) Demonstrate that technical risks have been identified and that solutions are in hand.

(2) Provide the technical data to support the ratification decision.

OTI: (1) To provide an indication of utility and worth to the user.

(2) Refine identified critical operational issues and seek new issues to be examined in subsequent testing.

(3) Provide the operational test data to support the independent evaluation furnished to ASARC II for the ratification decision milestone.

b. Task: Combined DT II/OT II (full-scale development phase)

Objectives:

DTII: (1) Provide technical data to support the transition of system development to either low-rate initial production or full-scale production.

(2) Demonstrate that engineering is essentially complete and that solutions to all significant design problems are in hand.

OTII: (1) Assess military utility, operational effectiveness and operational suitability in a realistic operational environment.

(2) Provide the operational test data to support the independent evaluation furnished to ASARC IIa concerning the decision to initiate either low-rate initial production or full-scale production.

c. Task: Combined DTIII and OTIII (full-scale development phase)

Objectives:

DTIII: (1) Validate production items.

(2) Confirm corrective actions for development problems disclosed in earlier testing.

(3) Conduct DT&E scheduled for, but not conducted in earlier development testing.

OTIII (1) Examination of initial production items to ensure that the system is operationally suitable; that all operationally critical issues have been resolved, and that all benefits and burdens of the system are identified.

(2) Provide the operational test data to support the independent evaluation concerning the ASARC III production decision.

General responsibilities in combined DT/OT

Developing Agency's Role	OTEA's Role
Provides test planning and test execution baseline.	Prepares separate Test Plan, coordinated with and using DT Test Plan as the baseline.
Provides test items such as developmental items and system prototypes or initial production models.	Provides Deputy Director for OT
Primary responsibility for the test program.	Ensures participation of representative user troops.
Provides Deputy Test Director for DT.	Utilizes data from DT, OT and other sources for evaluation.
Coordinates test resources in the Coordinated Test Program.	Coordinate test resources in the Coordinated Test Program.
Prepares separate DT report. Distributes report in accordance with Table 2 of AR 70-10.	Provides separate OT report.
	Provides independent evaluation to the decision body for the appropriate decision milestone.
	Distributes reports in accordance with Table 2 of AR 70-10.

Combat Developer's Role

Provides proposed concepts of employment (doctrine, tactics, technique, organization, logistical procedures and training).

Defines threat and environment.

Furnishes scenarios

Ensures user's interests are represented.

8. OT&E DEFINITION AND OBJECTIVES IN THE DEPARTMENT OF THE ARMY

A review of Army Regulations concerned with operational test and evaluation fails to reveal a definition for operational test and evaluation per se, or for follow-on operational test and evaluation.

The Army definitions and objectives concerning OT&E and the sources of definitions and objectives are provided as follows:

a. "Operational Testing - Testing of material systems which is conducted by an organization independent of the developer. OT is accomplished with representative user troops in as realistic an operational environment as possible to provide data to assess

(1) Military utility, operational effectiveness, and operational suitability (including compatibility, interoperability, reliability, availability, maintainability, and integrated logistics support and training requirements of new systems).

(2) From the user's viewpoint, the new system's desirability, considering systems already available and the operational benefits/burdens associated wit' the new system.

(3) The need for modifications to the system.

(4) The adequacy of doctrine, organization, operating techniques, tactics and training for its employment and the system for its maintenance support." AR 10-4: "U. S. Army Operational Test and Evaluation Agency"

b. "Operational Evaluation - The test and analysis of a specific end item or system, so far as practicable under service operating conditions, in order to determine if quantity production is warranted considering a. the increase in military effectiveness to be gained, and b. its effectiveness as compared with currently available items or systems, consideration being given to: (1) personnel capabilities to maintain and operate the equipment; (2) size, weight and location considerations; and (3) enemy capabilities in the field."

AR 310-25: "Dictionary of United States Army Terms"

c. OT&E Objectives -

(1) "Operational testing will be conducted to determine if the system is operationally suitable from a doctrinal, organizational, and tactical point of view and to collect performance and reliability, availability, and maintainability (RAM) data for the equipment when in the hands of troops.

Annex A of DAFD-SPY: "Letter of Instructions (LOI) for Implementing the New Material Acquisition Guidelines" dated 23 August 1972

(2) "(Operational) test and evaluation shall be conducted throughout the material acquisition process to assist in progressively eliminating acquisition risks and in assessing military worth .... The goal is to accomplish required testing of important characteristics before each decision point and within a maximum of six months from start of such testing."

Annex H of DAFD-SDY "Letter of Instruction (LOI) for Implementing the New Material Acquisition Guidelines" dated 23 August 1972

9. THE ARMY ORGANIZATION

A summary view of the organization of the Army is provided (Ref Figure C-5) followed by a more detailed description of the Army organization for the conduct of user testing. Also included are the identification of major participants and significant documentation related to the material acquisition process.

Summary view. The Secretary of the Army (SA), as civilian head of the Department of the Army (DA), is directly responsible to the Secretary of Defense for all organizational, administrative, and operational affairs of the DA. The principal assistants to the SA are the Undersecretary and Assistant Secretaries of the Army. The Army Staff is the military staff of the Department of the Army. The Chief of Staff, Army acts as agent of the SA and is head of the Army Staff. The Army Staff consists of the General Staff, the Special Staff, and the Personal Staff.

As the senior Army officer and head of the Army Staff, the Chief of Staff is the principal military advisor to the SA. He is responsible for the planning, developing, execution, review, and analysis of all Army programs and activities. A direct line of command authority extends from the Secretary of the Army to the Chief of Staff, Army.

The General Staff, under direction of the Chief of Staff, renders professional advice and assistance to the Secretary of the Army and the Assistant Secretaries of the Army, providing broad guidance, basic policies and plans. The General Staff also assists the SA in the preparation and issuance of directives to implement plans and policies and is responsible for the supervision and execution of such directives.

The Special Staff, under the Chief of Staff consists of the military heads of specialized functions which are essentially in support of the General Staff operation, such as the Adjutant General, the Chief of Engineering, the Surgeon General, etc.

The Chief of Staff, Army exercises control over the Major Army Commands and all military elements of the Department of the Army such as the material developers, the combat developers, and the active and reserve Army combat and support units.

Major Army Commands include Forces Command, U.S. Army, Alaska; U.S. Army, Europe; the U.S. Army Materiel Command, and others.

Organization for User Testing. Within the Department of the Army three principal organizations exist to perform the function of user testing. These organizations are:

a. OTEA - An Agency under the Chief of Staff, Army charged with the responsibility for OT&E of major and selected nonmajor systems, for the management of all OT&E, and for the coordination of FDTE and joint testing.

b. CDEC - (Combat Development Experimentation Command) - A subordinate command of Training and Doctrine Command (a Major Army Command), responsible for the development of concepts and for the conduct of field experiments involving both OT&E and FDTE. Normally, the OT&E and FDTE assigned to CDEC involves small scope, highly instrumented, objective testing.

c. MASSTER - (Modern Army Selected Systems Test Evaluation and Review) - A subordinate command of Forces Command (a Major Army Command), charged with planning and conducting OT&E and FDTE and testing concepts involving doctrine, material, organization and training. Normally the OT&E and FDTE assigned to MASSTER are the larger scope, less instrumented, subjective-type tests.

The Coordination of the User Testing Effort. The coordination of all user testing effort is accomplished through a mechanism called the Test Schedule and Review Committee (TSARC). The TSARC reviews and recommends for approval the proposed Army Five Year Test Program which is the tasking document to OTEA, CDEC, MASSTER, and all other commands and agencies for the conduct of user testing. The

TSARC is composed of:

CDR OTEA (Chairman), and general officer representative of:

TRADOC	CDEC
FORSCOM	TECOM
AMC	Other Army General/Special Staff, as required.
DA Staff (R&D, DCSOPS, DCSLOG, Comptroller)	Army Commands as appropriate to the level of involvement.
MASSTER	

Functions. The TSARC functions are:

a. Convenes twice each year (June, December) regularly. On call, as necessary to resolve critical problems.

b. Reviews the proposed Army Five Year Test Program (FYTP) (prepared by OTEA for the CSA) against the needs of test data to support specific decision points. (Decision of TSARC must be unanimous or are resolved by Vice Chief of Staff of the Army.)

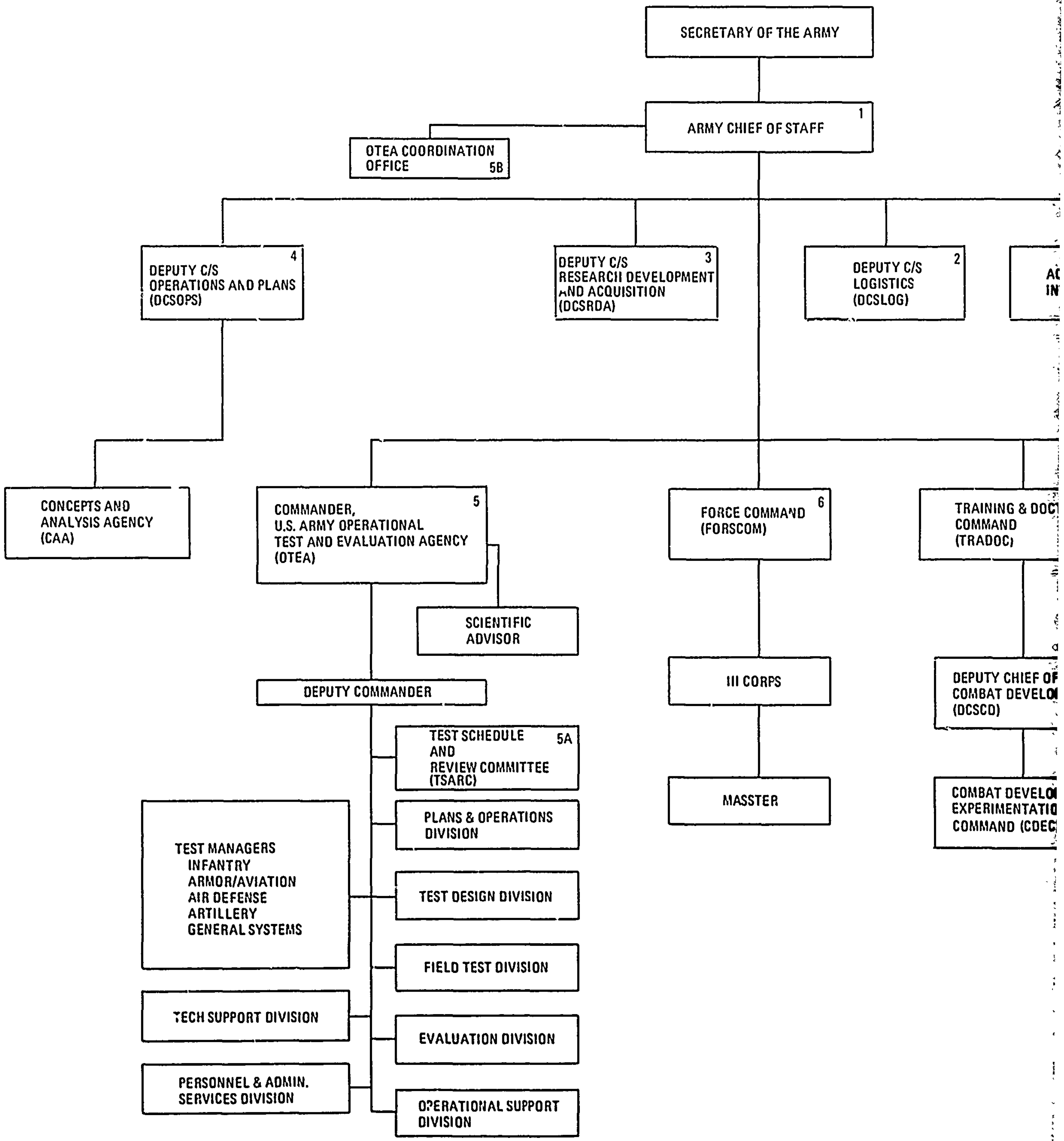
c. Recommends approval of FYTP.

FYTP. The Army Five Year Test Program (FYTP) contains Outline Test Plans (OTP's) for OT&E, proposals for FDTE, and Army requirements in the DOD program for joint testing. OTP's for major and selected nonmajor systems are prepared by OTEA. OTP's for other nonmajor OT&E are prepared by the command responsible for OT&E as designated by DA, usually TRADOC.

The FYTP contains personnel, material, facilities requirements (who, where, when, how) of user testing. It is directive for the current fiscal year +1 and is approved for planning and budget guidance for the more distant three years.

Responsibilities in Material Acquisition and Testing. The following identifies principal participants in Army material acquisition and testing and provides a summary description of the roles of each: (refer to like-numbered blocks in Figure C-5 "Army Organization for Material Acquisition and Test.")

Block Number	Title	Functional Description
1	Chief of Staff, Army	a. Insures adequacy of the Army's overall material acquisition and test program b. Provides policy direction and guidance in accordance with policies of SA.
2.	Deputy Chief of Staff Logistics (DCSLOG)	Primary Staff responsibility for overall policy for improvement of material (deployment phase) including related OT.
3	Deputy Chief of Staff for Research, Development and Acquisition (DCSRDA)	a. Establishes policy for Army R&D b. Plans, coordinates, supervises all Army research, development, test and evaluation. c. Establishes research and development objectives. d. Coordinates with DCSOPS to approve/disapprove ROC's for nonmajor systems. e. Exercises General Staff management for development of nonmajor systems (includes program reviews and controls.) f. Appoints using command or agency to coordinate with material developer on the development and test of nonmajor systems.


SECRETARY OF THE ARMY
ARMY CHIEF OF STAFF 1
OTEA COORDINATION OFFICE 5B
DEPUTY C/S OPERATIONS AND PLANS (DCSOPS) 4
DEPUTY C/S RESEARCH DEVELOPMENT AND ACQUISITION (DCSRDA) 3
DEPUTY C/S LOGISTICS (DCSLOG) 2
CONCEPTS AND ANALYSIS AGENCY (CAA)
COMMANDER, U.S. ARMY OPERATIONAL TEST AND EVALUATION AGENCY (OTEA) 5
FORCE COMMAND (FORSCOM) 6
SCIENTIFIC ADVISOR
DEPUTY COMMANDER
III CORPS
TEST SCHEDULE AND REVIEW COMMITTEE (TSARC) 5A
MASSTER
PLANS & OPERATIONS DIVISION
TEST MANAGERS INFANTRY ARMOR/AVIATION AIR DEFENSE ARTILLERY GENERAL SYSTEMS
TEST DESIGN DIVISION
FIELD TEST DIVISION
TECH SUPPORT DIVISION
EVALUATION DIVISION
PERSONNEL & ADMIN. SERVICES DIVISION
OPERATIONAL SUPPORT DIVISION

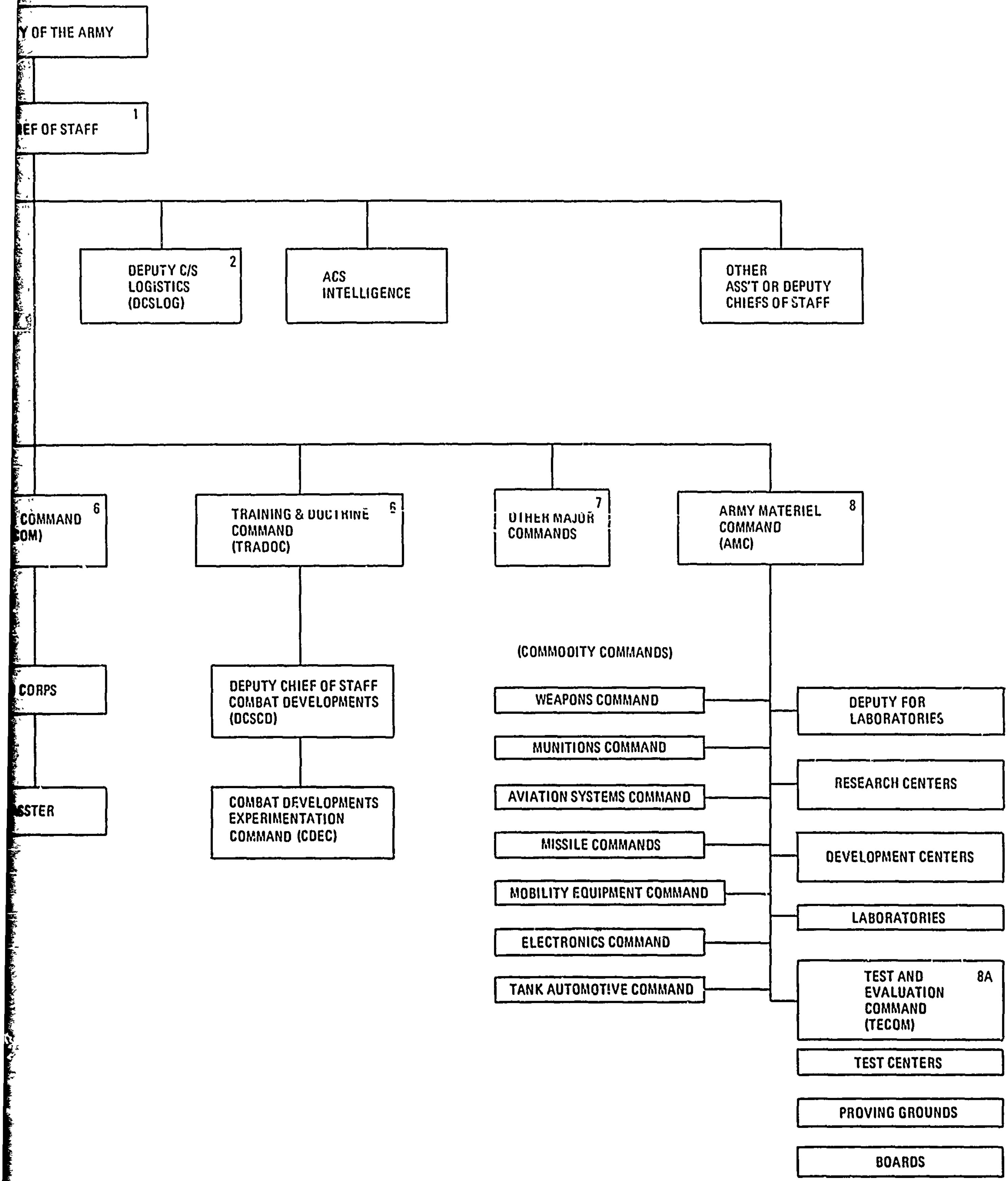


Figure C-5. Army organization for material acquisition and test

Block Number	Title	Functional Description
		g. Staff responsibility for the Coordinated Test Program (CTP) which includes both development testing and operational testing. h. Staff management responsibility for nonmajor systems, including adequacy of testing and level of program review.
4	Deputy Chief of Staff for Operations and Plans (DCSOPS)	a. Exercises General Staff responsibility for: (1) Development of broad force requirement. (2) Development of strategic concepts, plans, estimates. b. Issues appropriate guidance for R&D and Combat Development Programs through the "Basic Army Strategic Estimate (BASE) and the Army Strategic Plan (ASP). c. Approves ROCs. d. Primary Staff responsibility for monitoring the provision of troop resources for OT:
5	Operational Test and Evaluation Agency (OTEA)	a. Responsible to CSA for management of all user testing. b. Determines when, where, how and by whom OT will be accomplished. c. Prepares the OT portion of the Coordinated Test Program of the Development Plan.

Block Number	Title	Functional Description
		d. Plans, programs, budgets, directs, reports, and evaluates OT on all major and selected nonmajor systems.
		e. Provides inventory control, identifies requirements and provides assistance in test instrumentation for user testing.
		f. Coordinates planning, programming, budgeting, and scheduling for all OT.
		g. Establishes requirements, provides guidance for planning and conduct of OT on nonmajor systems.
		h. Provides the chairman for the Test Schedule and Review Committee (TSARC).
		i. Prepares and publishes the Army Five Year Test Program (FYTP).
		j. Functions as initial point of contact and coordinator for the OT portion of all combined DT/IOT and for Army participation in joint user testing.
5A	Test Schedule and Review Committee (TSARC)	a. Review and approve the OT, FDTE, and Joint Testing proposed for the Army Five Year Test Plan (FYTP).
		b. Unanimously approve the assignments of responsibility for conduct and support of user testing contained in the FYTP.

Block Number	Title	Functional Description
5B	OTEA Coordination Office	a. Functions as focal point organization at HQDA to keep DA and OSD fully informed on the Army's OT needs and accomplishments. b. Provides interface between CSA and OTEA for OTE matters.
6	Combat Developer	a. Formulates material objectives and requirements relating to employment of Army forces in a theater of operations for systems in development and test. b. Maintain cognizance of development program. c. Participate in the preparation of the Development Plan, including the Coordinated Test Program. d. Provide representation to management reviews of nonmajor systems. e. Participate in test planning, monitor conduct of tests, and review test reports. f. Provide appropriate comments to DA.
7	CONUS and Overseas Commands	a. Provide supplementary support to OT&E as required by the FYTP. b. Provide user's inputs into the system under development.

Block Number	Title	Functional Operation
		c. Pursuant to approved test plan assignments in the FYTP, provide administrative, logistic, and military personnel/units for OT.
8	Material Developer	a. Conduct research, development, test and evaluation (RDT&E). b. Translate operational requirements into hardware system. c. Conduct maintenance support planning. d. Manages technology base development effort. e. Designates the Project Manager for major Army systems, or otherwise assigns responsibilities for lesser acquisition programs in order to: (1) Coordinate and execute a T&E program responsive to DT&E/IOT&E questions and issues. (2) Prepare a Coordinated Test Program (CTP) early in the acquisition process to coordinate and integrate DT and OT. f. Establish test support requirements for material acquisition.

Block Number	Title	Functional Operation
		g. Preparation, coordination, distribution and maintenance of the Development Plan (DP).
		h. Provide T&E technical support.
		i. Coordinate with OTEA on T&E planning and support.
		j. Furnish material and technical support, including spare parts and special test equipment.
		k. Provide necessary funding to activities supporting DT.
8A	Test Command (TECOM)	a. Coordinate with appropriate commands/agencies within AMC and with OTEA and others to define test requirements and objectives.
		b. Provide membership to Special Task Force and provide inputs to the DCP and the Coordinated Test Program.
		c. Plan, conduct, report DT&E.
		d. Integrate OT into DT testing for combined DT&E/IOTE testing.
		e. Provide DT&E results through AMC chain of commands to ASARC/DSARC at decision milestones in the acquisition process.

Block Number	Title	Functional Description
(not shown)	Special Task Force	Appointed in the conceptual phase by Army Chief of Staff to accomplish the following: (1) Provide necessary data concerning a proposed major system to support a DA/OSD decision to initiate development. (2) Produce a draft Development Concept Paper (DCP), a Concept Formulation Package (CFP), and a Final Report which includes a plan for test and evaluation.
(not shown)	Army Systems Acquisition Review Council (ASARC)	Review major system development programs designated by HQDA to obtain a HQDA decision at key decision milestones. (i.e. the decisions to enter the validation phase (ASARC I); the full scale development phase (ASARC II); and the production phase (ASARC III) in the system acquisition process.

10. SIGNIFICANT DOCUMENTS IN THE ACQUISITION PROCESS

The significant documents peculiar to the Army requirements/acquisition process and the normal order of introduction of such documents into a given acquisition process is presented as follows:

a. Operational Capabilities Objectives (OCO) - The OCO is a document which sets forth the operational capabilities required by the Army within a future period of ten or more years. Such requirements are stated in terms of general capabilities. Their inclusion is the OCO serves to provide direction and guidance to Army research and does not constitute approval for material development actions.

b. Catalogue of Approved Requirements Documents (CARDS)

The CARDS is a requirements document which is updated and issued quarterly by the DCSOPS. The CARDS consist of three sections, identified as follows:

(1) Section 1 - Contains the Operational Capabilities Objective (OCO) described above.

(2) Section 2 - This section is titled "Approved Material Requirements Document" and is a compilation of ROCs which have been approved and are either dormant or have resulted in development actions not yet completed. An approved ROC for which funding is not established will be carried within Section 2 for two years, after which the ROC is deleted. ROC's within this section which have resulted in funding and development actions are identified by:

(a) Title of system

(b) Material developer

(c) Combat developer

(d) Decision level (i.e. DSARC/ASARC/Army in-process review (IPR).

(3) Section 3 - Entitled "Deleted Requirements", this section contains the listing of ROCs which were previously contained in Section 2, but which were deleted after a period of 2 years due to a lack of funding resulting from either the demands of higher priority requirements, a change of threat, or other reasons.

c. Concept Formulation Package (CFP) - The CFP, as discussed earlier, is one of the products of the Special Task Force subsequent to the ROC approval for a major system. On a case by case basis, and as directed by DA, a CFP may be prescribed also for lesser systems in which case

its preparation is the responsibility of the material developer. The CFP contains the supporting documentation for the draft DCP and consists of the following:

(1) Trade-off Determination (TOD)

(2) Trade-off Analysis (TOA)

(3) Best Technical Approach (BTA)

(4) Cost and Operational Effectiveness Analysis (COEA)

d. Final Report - Another product of the Special Task Force is the Final Report. This document is provided to the Program Manager upon his assumption of program responsibility (normally at or immediately after the approved program decision at ASARC I/DSARC I). The Final Report is the basis for and provides significant input to the Development Plan (DP). The Final Report contains:

(1) Systems summary

(2) Systems requirement and analysis

(3) Alternatives in approach to system development, and relationships to other systems.

(4) Plan for system development

(5) Technical portion of the RFP

(6) Financial and procurement plan

(7) Plan for test and evaluation (provided by the material developer and OTEA).

e. Development Plan (DP) - The DP corresponds to the Program Management Plan in Air Force development programs. It provides appropriate analysis of technical options and the life cycle plans for development, production, training and support of material items, records program decisions, and contains the user's requirements. The material developer,

represented by the Program Manager, is responsible for the DP which is used for both major and nonmajor programs. Elements of the DP are:

(1) System summary

(2) System requirements and analysis

(3) Plans for systems development

(4) Coordinated Test Program

(5) Plan for personnel and training requirements

(6) Logistic support plan.

f. Coordinated Test Program (CTP). The CTP (section IV of the Development Plan) contains the overall coordinated plan for testing in the entire acquisition cycle. The material developer's tester prepares the DT&E portions of the CTP (chapters 1 and portions of chapters 3 and 4). The command/agency assigned OT&E responsibilities, which for major and selected nonmajor systems is OTEA, prepares the OT&E portion of the CTP (chapter 2 and portions of chapter 3 and 4). The subsequent development of Outline Test Plans is based on and become an attachment both to the CTP and to the schedules and details for tests contained in the Army Five Year Test Program. (FYTP).

11. IMPACT OF DOD DIRECTIVES 5000.1 AND 5000.3

From 1970 through 1973, significant changes were effected in the DOD approach to acquisition, test, and evaluation of systems and material. These changes were initially the subject of memoranda to the Secretaries of the Military Components and were later formally promulgated as DOD Directives. The first of these Directives was DOD Directive 5000.1, issued in 1971, which established the DOD policy and provided guidance for the acquisition of major defense systems. This directive defined major defense systems, identified responsibilities in acquisition assigned to the military departments of the DOD and those specific decisions and responsibilities to be

exercised by the Secretary of Defense; established the Program Manager's role, provided for the conduct of development programs in phases identified as conceptual, validation, full scale development, production, and deployment; and established guiding program considerations in the areas of system needs, cost, logistic support, the reduction of risks and management control. Further, the directive specified that test and evaluation would commence early in the acquisition process and would provide the data on which large-scale production commitments would be based.

In January 1973, DOD Directive 5000.3; "Test and Evaluation" was issued. This directive applied certain management principles to nonmajor as well as major acquisition programs, and most significantly, directed that within each DOD component (Army, Navy, Air Force) "there will be one major field agency separate and distinct from the developing/processing command which will be responsible for OT&E."

For the Army, the DOD Memoranda and the formal directives (DOD Directives 5000.1 and 5000.3) which followed had great impact, requiring changes in Army procedure, policy, and organization; the more significant of which were:

a. The separation of operational testing from development testing under the material developer, and the relocation of the OT&E functions within Combat Development Command (CDC), an already existing command within the Army structure.

b. The subsequent disestablishment of CDC and the establishing of the U.S. Army Operational Test and Evaluation Agency, a Field Agency reporting directly to the Chief of Staff, Army for the management of all user testing (OT&E, FDT&E, and Joint Testing).

c. The realignment of the acquisition phases recognized by the Army to conform to the acquisition phases and decision milestones set forth in DOD Directive 5000.1.

d. Changes in acquisition program management, documentation and testing in order to satisfy DOD requirements.

Army Implementation of DOD Policy for Acquisition and Test. The Army directives listed below implement the DOD policies set forth in DOD Directives 5000.1 and 5000.3 and within the Department of the Army, establish policy, provide guidance and assign responsibility concerning system/equipment acquisition and test and evaluation:

a. Department of the Army Memorandum for Major Army Commanders and Heads of Army Staff Agencies, dated 20 June 72; subject: "Material Acquisition Guidelines".

b. Department of the Army Letter of Instructions (LOI) for Implementing the New Material Acquisition Guidelines, dated 23 Aug. 72.

c. Army Regulation (AR) 1000-1, dated 30 June 72; subject: "Basic Policies For Systems Acquisition By The Department of The Army".

d. Army Regulation (AR) 70-10, dated 26 Oct 73; subject: "Test and Evaluation During Development and Acquisition of Material".

e. Army Regulation (AR) 10-4, dated 15 Jan 74; subject: "U.S. Army Operational Test and Evaluation Agency".

f. Department of the Army Letter of Instruction (LOI) for Management of Joint User Testing Programs, dated 28 Feb. 74.

12. PRINCIPAL ARMY TEST AND EVALUATION FACILITIES

Arctic Test Center, Ft. Greeley, Alaska
Tropic Test Center, Ft. Clayton, Canal Zone
Proving Grounds:

Aberdeen, Md.
Dugway, Utah (includes the Desert Test Center).
Electronics, Ft. Huachuca, Arizona
Jefferson, Indiana
White Sands Missile Range, New Mexico

Kwajalein Missile Range, Marshall Islands

Combat Development Experimentation Command (CDEC)

Modern Army Selected Systems Test Evaluation and Review (MASSTER)

Test Command (TECOM) Boards.

References:

AR 10-4 "US Army Operational Test and Evaluation Agency"
15 Jan 74

HQ DA Pamphlet No. 11-25 "Life Cycle Management Models For Army Systems"
(Draft) Dec. 72

AR 1000-1 "Basic Policies For Systems Acquisition By The Department of
the Army" 30 Jun 72

AR 70-10 "Test and Evaluation During Development and Acquisition of Material"
26 Oct. 73

DA Letter of Instructions (LOI) for Implementing the New Material Acquisition
Guidelines 23 Aug. 72

DA (DAFD) letter "Responsibilities for User Testing" 27 Feb. 73

DA Letter of Instructions (LOI) for Management of Joint User Testing Programs
28 Feb. 74

OTEA "Operational Test and Evaluation METHODOLOGY GUIDE", (Draft) Feb. 73

MASSTER "Test Officers Planning Manual" Feb 74

CDEC "CDEC Experimentation Manual" Nov. 73

Appendix D

NAVY OPERATIONAL TEST AND EVALUATION - AN OVERVIEW

1. INDEPENDENT TEST AGENCY FOR OT&E

OPTEVFOR. The Operational Test and Evaluation Force (OPTEVFOR), under the direct command of the Chief of Naval Operations (CNO) for the conduct of OT&E, is the Navy's independent test agency for OT&E. The command headquarters, COMOPTEVFOR, is located ashore in the Camp Allen Annex of the Norfolk (Virginia) Naval Base.

2. BACKGROUND

The History of the Navy's Independent Test Agency. OPTEVFOR traces its origin to the final months of World War II. During the Okinawa campaign, the surface forces of the Navy operating in support of the ground forces on Okinawa came under a sustained concentrated attack by Japanese kamikaze("divine wind")aircraft. The suicide attacks by these aircraft took a great toll in ships and personnel, and a means was urgently required to deal effectively with the threat. In July of 1945, the Composite Task Force, U. S. Atlantic Fleet was formed and tasked with the development and evaluation of methods to combat such attacks. This was the origin of the Navy's independent operational test agency.

In the years that followed, changes were made to the name of the command and to the assigned mission and tasks to provide for a wider scope of responsibilities in test and evaluation. Changes were also made in the organizational structure which expanded test and evaluation capabilities, formed a similarly structured command for test and evaluation activity within the Pacific Fleet, and located the command ashore. The significant dates in the history of the Operational Test and Evaluation Force are.

July 1945 - Formation of Composite Task Force, U. S. Atlantic Fleet

1947 - Redesignated as "Operational Development Force, U. S. Atlantic Fleet"

1949 - Location of the command ashore at the Norfolk, (Virginia) Naval Base

1959 - Revision of mission to include wider OT&E responsibilities. Renamed "Operational Test and Evaluation Force."

Mission and Tasks. The mission and functions of OPTEVFOR are assigned by the Chief of Naval Operations and are as follows:

a. Mission: "To operationally test and evaluate specific weapon systems, ships, aircraft and equipments, including procedures and tactics, where required; and, when directed by CNO, assist development agencies in the accomplishment of necessary Development Test and Evaluation."

b. Specific Tasks: Specific tasks are assigned to COMOPTEVFOR by CNO which are summarized as:

(1) Function as an independent test agency for OT&E under the command of CNO. Serve as principal advisor to the CNO for all Department of the Navy OT&E matters.

(2) Present results of OT&E to the DSARC III (production decision) review and to other reviews as directed by CNO.

(3) Conduct operational tests on weapon systems including ships and aircraft. Evaluate operational effectiveness, suitability, and capability, reporting results to CNO.

(4) Assist development agencies in DT&E, including Fleet support, as required, reporting to CNO the results of such assists and an assessment of the system tested.

(5) Review and evaluate the T&E planning for new weapons systems to address and resolve critical issues and report findings to the CNO.

(6) Monitor and report other T&E efforts as directed by CNO.

c. Assigned Functions:

(1) Early DT&E Involvement: The CNO has also formally assigned functions to COMOPTEVFOR which ensure early participation in the development test and evaluation (DT&E) of an emergent system. During the first

three phases of the acquisition process (Conceptual, Validation and Full-Scale Development Phases), COMOPTEVFOR, as directed by CNO, serves as the coordinator between the developing agency and fleet units in arranging for services and facilities (including an operational environment) required by the developed for research, investigatory and development projects, and for development testing. In these instances, COMOPTEVFOR also functions as an observer and evaluator, reporting to CNO on the developing agency's project. Early involvement is assured by the following functions assigned by CNO:

(a) (Validation Phase) Providing to the developing agency operational planning assistance in determining the testing required, anticipating future testing requirements, and defining critical operational issues.

(b) (Validation Phase) Participation in testing planning, independently observing selected tests and demonstrations conducted by the developing agency, and examining test results. Submission of test assessment to CNO and the developing agency to include comments and recommendations concerning future operational suitability of the system, progress to date, and operational issues requiring further examination. (note: this function introduces initial OT&E, combined with DT&E, in the Validation Phase).

(c) (Full-Scale Development Phase) Independently observing selected developmental tests and reviewing the data from an operational viewpoint in order to verify readiness for operational testing.

(2) Initial OT&E Functions: As noted above, the formally assigned functions of COMOPTEVFOR specify that COMOPTEVFOR will commence the

initial OT&E of a system in the Validation Phase. Additional IOT&E functions assigned are:

(a) (Full-Scale Development Phase) Planning and conducting operational tests of specific equipments or systems

(b) (Full-Scale Development Phase) Provide IOT&E results to CNO prior to the major production decision.

(3) Follow-on OT&E (FOT&E) functions: The following FOT&E functions are formally assigned to COMOPTEVFOR:

(a) (Production and Deployment Phases) Plan and conduct FOT&E to develop optimum procedures or tactics.

(b) (Deployment Phase) Conduct further OT&E as assigned by CNO to verify correction of deficiencies and to validate system performance.

Command Relationships: COMOPTEVFOR is under the direct command of the CNO for the conduct of OT&E. Other command relationships are:

a. CNO provides:

Policy direction

Financial support

Technical and procedural guidance

b. CINCLANTFLT provides:

Administrative support for COMOPTEVFOR HQs Staff and logistic services for Atlantic Fleet operations

c. CINCPACFLT/CINCUSNAVEUR provide:

Logistics services for fleet operation conducted by COMOPTEVFOR in respective areas.

d. CONCLANTFLT/CINCPACFLT/CINCUSNAVEUR, as appropriate, exercises operational control over Fleet units assigned to COMOPTEVFOR.

3. OPTEVFOR INVOLVEMENT IN THE ACQUISITION CYCLE

From the discussion of mission, task and functions assigned by CNO to OPTEVFOR (paragraph 2; Mission and Tasks), it may readily be determined that not only is OPTEVFOR the independent Navy agency for operational test and evaluation, but that OPTEVFOR also performs the function of an interface between the research or development agency and the fleet for all RDT&E projects assigned by CNO involving the participation or assistance of operational units. The full scope of such participation by OPTEVFOR ranges from basic research projects not necessarily related to any specific development program to projects which are program oriented. In an acquisition cycle of a given system or equipment, OPTEVFOR may be involved in any or all of the following test and evaluation projects:

	Symbol
Fleet Research Investigation	R
Development Assist	D
Operational Assist	X
Technical Evaluation Project (TECHEVAL)	T
Operational Evaluation (OPEVAL)	O
Fleet Operational Appraisal	F

Project Identification: The projects listed above are assigned by CNO and are identified by a long and a short title. The long title is a concise phrase descriptive of the project. The short title is a corresponding letter/number which is to be included in all written communications concerning the project. The short title consists of the following:

The letter symbol, shown above for the type of project followed by a slash bar (/) and a second letter symbol denoting shipboard applications (/s); airborne application (/v); or surface and/or sub-surface application (/c).

A number which indicates the sequential order of assignment by CNO. (Example of the foregoing: o/s 54- An Operational Evaluation applying to shipborne equipment, assigned number 54 in the CNO series.)

Project Definitions. Projects assigned by CNO are categorized as:

a. Fleet Research Investigation - An examination by the developing agency, not necessarily program oriented, or natural or special phenomena in operational environment conducted for the collection of research data and requiring the assistance of operating forces of the Navy.

b. Development Assist - A CNO project for providing fleet services and support to the developing agency during the Conceptual Phase for the tests needed in gathering data to determine the direction in which development should proceed. Development Assists may also relate to material improvements of equipment already operationally deployed (e.g., proposed SHIP-ALTS, ORDALTS).

c. Operational Assist - A Validation Phase project assigned by CNO in which the developing agency is provided services and assistance of Navy operational units for the purposes of establishing confidence in program worth and the development effort, including the development test and evaluation (DT&E) data required to support the review and decision process (i.e., DSARC II) prior to the initiation of the Full-Scale Development Phase. An Operational Assist may also apply to certain material improvement programs.

d. Technical Evaluation Project (TECHEVAL) - A CNO project assigned and conducted during the Full-Scale Development Phase for the purpose of investigating systems or equipment and collecting data to permit the developing agency to determine whether the system or equipment is technically

acceptable, meets design and performance specifications, and is technically suitable for Operational Evaluation (OPEVAL). For aircraft or missile development programs, the TECHEVAL will include the Navy Preliminary Evaluation (NPE) or the Navy Technical Evaluation (NTE). TECHEVALS may apply, also, to material improvement programs including conversions, major modifications and modernization.

(1) Navy Preliminary Evaluation (NPE) - An early assessment in the development readiness for trials by the Board of Inspection and Survey (BIS). An NPE is usually conducted at the developing contractor's plant; however, a portion of the evaluation may utilize certain military facilities, such as bombing, gunnery and missile ranges.

(2) Navy Technical Evaluation (NTE) - NTEs are Navy missile evaluations conducted as AIRTASKS by Naval Air Systems Command or Naval Ordnance Systems Command for the purpose of assessing the missile's potential for satisfying the operational requirement and design objectives. As a general rule, NTEs are completed just prior to Operational Evaluation and the deployment of the missile to the operational forces.

e. Operational Evaluation (OPEVAL) - A CNO project assigned during the Full-Scale Development Phase to COMOPTEVFOR for execution, assisted as necessary by the development agency, the objectives of which is to conduct IOT&E in order to satisfy test and evaluation data requirements for the program review and production decision. The conclusion of the OPEVAL completes the IOT&E process unless further tests are directed by the program review authority as a requisite for a production decision.

f. Fleet Operational Appraisal - The Fleet Operational Appraisal is a CNO project assigned during the Production or Deployment Phase to COMOPTEVFOR, or to commands recommended by COMOPTEVFOR,

for the conduct of FOT&E. When not assigned primary responsibility for the conduct of the project, OPTEVFOR provides assistance in planning and data analysis as requested.

4. COMBINED DT&E/IOT&E

DOD Directive 5000.3, section IV.3C states that DT&E and OT&E should be conducted separately. However, DOD policy does permit the combining of the two areas of test and evaluation where separation would cause delay involving unacceptable military risk, or would cause an unacceptable increase in acquisition cost.

Governing Navy directives recognize that the primary objective of the DT&E program is the generation of essential, valid data upon which to base design decisions during the development phases. The directives recognize also the essential consideration that the objectives of the DT&E program not be overburdened by secondary objectives. Nevertheless, the Navy philosophy concerning test and evaluation is deeply rooted in the belief that in addition to the achievement of design characteristics, the achievement of specific operational characteristics is a major responsibility of the developing activity and that development test plans (DT&E) must be structured and executed in such a manner as to ensure the generation of data required for the assessment of operational effectiveness and suitability (operational evaluation). In order that DT&E properly address such operational concerns, OPTEVFOR is assigned certain specific functions during system development which assist and supplement the role of the system developer in the generation of critical questions and issues, the formulation of test objectives, and in test planning and execution. The Navy's premise, therefore, is that DT&E and OT&E are not separable in a clean, clear fashion.

For purposes of this discussion of the combined Navy DT&E/IOT&E in a major system acquisition cycle, there are three test and evaluation projects in which the

Developing Agency and OPTEVFOR normally conduct an integrated effort. These combined DT&E/OT&E projects are:

a. The Development Assist Project

b. The Operational Assist Project

c. The Technical Evaluation Project

The matrices shown below depict the role of both the Developing Agency and OPTEVFOR in the combined DT&E/OT&E projects noted above.

(1) Task: DEVELOPMENT ASSIST PROJECTS (Projects conducted in the Conceptual Phase for the purpose of obtaining data required for development decisions and the request for a DSARC I program decision, or for system modification decisions).

DEVELOPING AGENCY'S ROLE	OPTEVFOR's ROLE
Primary responsibility for the project	Arranges for fleet services
Identifies fleet services required	Concurrence in test plan required in matters pertaining to fleet unit participation
Plans tests	Obtains fleet services, provides liasion
Coordinates test plan with OPTEVFOR	Monitors testing
Conducts test and evaluations	Independently reports to CNO with comments and recommendations, as appropriate
Reports to CNO on test and evaluation	Reports to CNO the fleet services provided for the task

(2) Task: OPERATIONAL ASSIST PROJECTS (A project conducted in the Validation Phase for the purpose of obtaining DT&E data to support the request for a DSARC II decision for full-scale development).

DEVELOPING AGENCY'S ROLE	OPTEVFOR's ROLE
Primary responsibility for the project	Participates in test planning
Identifies fleet services required	Arranges for fleet services
Plans tests, structured to provide data for assessment of suitability and effectiveness, as practicable	Participates in selected tests
Conducts tests (with participation of OPTEVFOR)	Participates in the inteprctation of test results
Interprets test results (with participation of OPTEVFOR)	Refines critical operational issues and T&E requirements contained in the DCP or comparable CNO document
Refines critical issues and questions in DCP or comparable CNO document, incorporating inputs from OPTEVFOR.	Reports "independent operational assessment" of the Operational Assist Project to CNO
Reports to CNO the results of Operational Assist Project test and evaluation	Recommends to CNO the operational issues and test requirements for the full-scale development phase.
Refines development issues and test requirements for full-scale development phase	

(3) Task: TECHNICAL EVALUATION (TECHEVAL) PROJECT: (A mandatory project in support of an Operational Evaluation (OPEVAL) Project, conducted during the full-scale development phase for the purpose of obtaining DT&E data to support both the following OPEVAL and the request for a DSARC III decision for production.)

DEVELOPING AGENCY'S ROLE	OPTEVFOR's ROLE
Primary responsibility for the project	Arranges for fleet services
Identifies fleet services required	Coordinates with Developing Agency in the refinement of the Test and Evaluation Master Plan (TEMP) and in planning the TECHEVAL tests
Provides prototypes, pilot production units, as available.	Identifies OPTEVFOR data requirements
Plans tests, to include the coordinated inputs of OPTEVFOR	Participates in selected tests
Conducts tests (with participation of OPTEVFOR)	Interprets test results
Inteprets test results	Reports "independent operational assessment" of the TECHEVAL to CNO.
Reports to CNO the results of the TECHEVAL	Refines operational issues to be addressed during OPEVAL.
Certifies the system is in a state of development readiness for the OPEVAL	

Sequence of Testing

In sequence in the acquisition cycle, the Operational Test and Evaluation (OPEVAL) Project follows the completion of the TECHEVAL which combines DT&E and IOT&E. The OPEVAL is concerned only with IOT&E issues. The planning and conduct of the OPEVAL tests, and the reporting of test and evaluation findings are the responsibility of OPTEVFOR. The completion of the OPEVAL marks the termination of the IOT&E effort by OPTEVFOR.

5. OT&E DEFINITION AND OBJECTIVES IN THE DEPARTMENT OF THE NAVY

Definition of Operational Test and Evaluation (OT&E). "Tests and evaluation participated in or performed by operational personnel focusing on operational effectiveness and suitability (including reliability, compatibility, interoperability, maintainability, and supportability). It also includes the development of optimum operational tactics for systems and equipment being developed for service use."

a. "Initial Operational Test and Evaluation (IOT&E): That T&E accomplished prior to the DSARC Milestone III or comparable CNO or CHNAVMAT first major production decision point to permit assessment of the operational effectiveness and suitability of a weapon system."

b. "Follow-On Operational Test and Evaluation (FOT&E): FOT&E is the continuing operational test and evaluation of a weapon system conducted in an operational environment by operational personnel using production systems for the purpose of verifying system performance; validating correction of deficiencies previously identified, and refining tactical employment doctrine and requirements for personnel and training. FOT&E may be initiated using pilot or pre-production systems which most closely resemble the production units until the latter items are available." OPNAVINST 3960.8

OT&E Objectives. "The objectives of the overall operational test and evaluation effort for any program is to aid in providing, at major decision points in the development and acquisition process, the best information possible at that point in time as to: the military utility of the prospective system; its expected operational effectiveness, operational suitability (including reliability, maintainability, simplicity, logis tic, and training requirements); need for modification; and the organization, doctrine and tactics for system deployment." SEC NAVINST 5000.1

6. THE NAVY ORGANIZATION

General: An understanding of the Navy's organization for test and evaluation requires a basic knowledge of the broader organization of the Navy. The Department of the Navy consists of the Navy Department, the Shore Establishment, and the Operating Forces. The Navy Department includes the Offices of the Secretary and Undersecretary of the Navy and associated staff offices, the four offices of Assistant Secretaries of the Navy and the management offices specifically assigned to each, the Chief of Naval Operations (CNO), the Commandant of the Marine Corps, and the Chief of Naval Material (CHNAVMAT). (Historically, the term "Navy Department" has meant the administrative and management offices of the Department of the Navy which are located at the seat of the Federal Government in Washington, D. C.) The Shore Establishment is comprised of the Systems Commands under CHNAVMAT and the various Offices, Bureaus, and Commands and shore activities of the Navy as well as the Corps Reserve and the Supporting Establishment of the Marine Corps. The Operating Forces of the Navy and Marine Corps are the forces afloat and the Fleet Marine Forces, respectively, and the commands associated with them, such as Commander in Chief, Atlantic Fleet (CINCLANFLT), and Commanding General, Fleet Marine Force (FMF), Atlantic.

A direct line of command authority extends from the Secretary of the Navy (SECNAV) to the senior line naval officer (THE CNO) and to the senior line Marine Corps officer (the Commandant of the Marine Corps). The CNO exercises control over CHNAVMAT, the Systems Commands of Air, Ordnance, Electronics, Ships, Supply and Facilities Engineering under CHNAVMAT, and the Bureaus, Offices, Commands and activities which, with the Systems Commands, comprise the Shore Establishment. The CNO also exercises direct control over the Operating Forces.

For the Marine Corps, the Commandant of the Marine Corps is the counterpart officer of CNO. He controls the Supporting Establishment, the Corps Reserve, and the Operating Forces of the Marine Corps.

Organizations for Test and Evaluation. Within the Department of the Navy, three separate organizations exist to perform the function of independent evaluation of systems and equipment. These organizations are Operational Test and Evaluation Force (OPTEVFOR), charged with the primary responsibility for Navy OT&E; the Board of Inspection and Survey (BIS), which conducts acceptance trials of ships and aircraft; and the Development Center, Marine Corps Development and Education Command (MCDEC), an organization with similar responsibilities and functions to OPTEVFOR, responsible for independent OT&E of systems and equipment for the Marine Corps. A review of the chart (Figure D-1) included herein, 'Navy Organization For Operational Test and Evaluation (OT&E) will provide additional perspective.

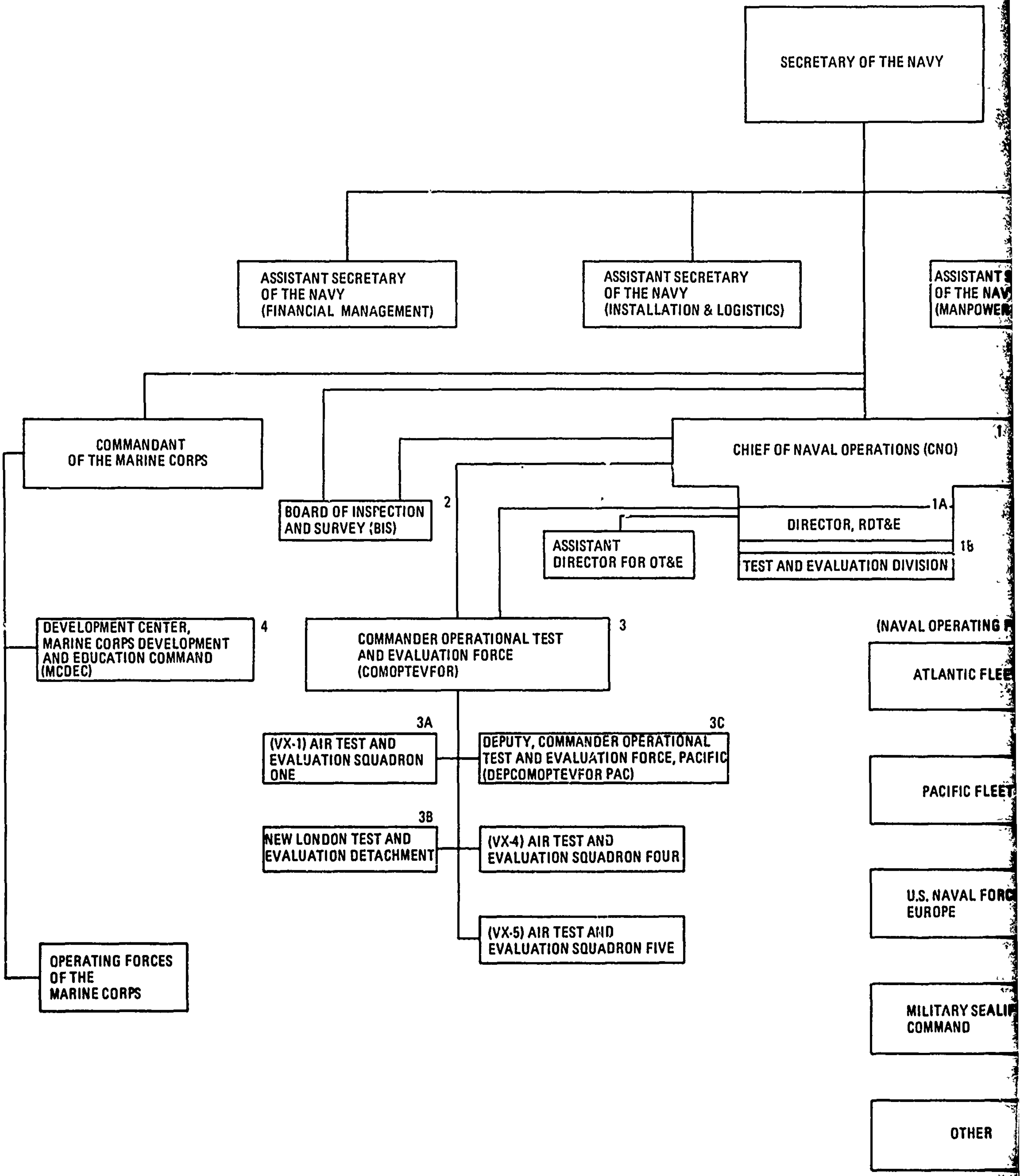
OPTEVFOR is the Navy's independent OT&E test agency, reporting directly to the CNO and to the Atlantic and Pacific Fleet Commanders. With resources and personnel in modest numbers on the east and west coasts of the United States, the command relies heavily on the resources and personnel of the Shore Establishment and Operating Forces for assistance in the prosecution of test and evaluation projects assigned. OPTEVFOR advises the CNO on the capability of new systems and equipment to satisfy Navy operational requirements as well as performance specifications and responds to the basic issue, "Does the system or equipment satisfy the operational requirement of the Fleet?"

The Board of Inspection and Survey (BIS), established during the administration of President Thomas Jefferson, is also an independent agency, reporting to the SECNAV through the CNO on the acceptability of ships and aircraft for service use. The Board is composed of a President and a small permanent staff and is assisted as required by technical agencies and personnel and resources from other organizations for the conduct of particular acceptance trials. The role of BIS is separate and distinct from the role of OPTEVFOR. The objective of the Board of Inspection of Survey is to determine whether the Navy (the buyer) got what it contracted for and whether the contractor (the seller) adequately fulfilled his guarantees. Thus, the roles of OPTEVFOR and BIS are different; however, the end result of the functions performed by each complements and strengthens the information available to CNO and SECNAV on which to base management decisions.

NAVY Organization Functional Descriptors For OT&E. Refer to like-numbered blocks in the chart titled "Navy Organization For Operational Test and Evaluation", Figure D-1.

Block Number	OT&E Functional Descriptor
1	Chief of Naval Operations: Insures adequacy of the Navy's overall test and evaluation program. Provides policy direction, technical and procedural guidance and financial support to COMOPTEVFOR in accordance with overall policies of SECNAV.
1A	Director, Research, Development Test and Evaluation: Serves as focal point within Office of CNO for T&E. Reviews proposed T&E objectives and requirements for fleet services to support T&E programs. Prioritizes all T&E programs and assigns to operating forces for prosecution. Receives, reviews, assesses all OT&E reports and functions as sole releasing authority of OT&E information to higher authority. Funding for T&E in the programming and budgetary system.

Block Number	OT&E Functional Descriptor
1A (cont)	Recommends CNO decision at key decision milestones for the further development and acquisition of less than major systems for which CNO is assigned responsibilities.
1B	Test and Evaluation Division: Controls the planning, conduct and reporting of all air, surface, and undersea/strategic test and evaluation. Coordinates assignment of OT&E projects to COMOPTEVFOR.
1C	Assistant-Director For OT&E: Provides staff assistance to the Director, RDT&E in all OT&E matters.
2	Board of Inspection and Survey: Conducts service acceptance trials of new ships and new model aircraft. Reports results of tests to the CNO and to the SECNAV.
3	Commander, Operational Test and Evaluation Force (COMOPTEVFOR): Functions as the Navy's independent test agency for OT&E Serves as principal advisor to CNO for all matters pertaining to Navy OT&E.


SECRETARY OF THE NAVY
ASSISTANT SECRETARY OF THE NAVY (FINANCIAL MANAGEMENT)
ASSISTANT SECRETARY OF THE NAVY (INSTALLATION & LOGISTICS)
ASSISTANT S OF THE NAV (MANPOWER
COMMANDANT OF THE MARINE CORPS
CHIEF OF NAVAL OPERATIONS (CNO)
1
BOARD OF INSPECTION AND SURVEY (BIS)
2
1A
DIRECTOR, RDT&E
ASSISTANT DIRECTOR FOR OT&E
1B
TEST AND EVALUATION DIVISION
DEVELOPMENT CENTER, MARINE CORPS DEVELOPMENT AND EDUCATION COMMAND (MCDEC)
4
COMMANDER OPERATIONAL TEST AND EVALUATION FORCE (COMOPTEVFOR)
3
(NAVAL OPERATING F
ATLANTIC FLEE
3A
(VX-1) AIR TEST AND EVALUATION SQUADRON ONE
3C
DEPUTY, COMMANDER OPERATIONAL TEST AND EVALUATION FORCE, PACIFIC (DEPCOMOPTEVFOR PAC)
PACIFIC FLEET
3B
NEW LONDON TEST AND EVALUATION DETACHMENT
(VX-4) AIR TEST AND EVALUATION SQUADRON FOUR
U.S. NAVAL FORC EUROPE
(VX-5) AIR TEST AND EVALUATION SQUADRON FIVE
OPERATING FORCES OF THE MARINE CORPS
MILITARY SEALIF COMMAND
OTHER

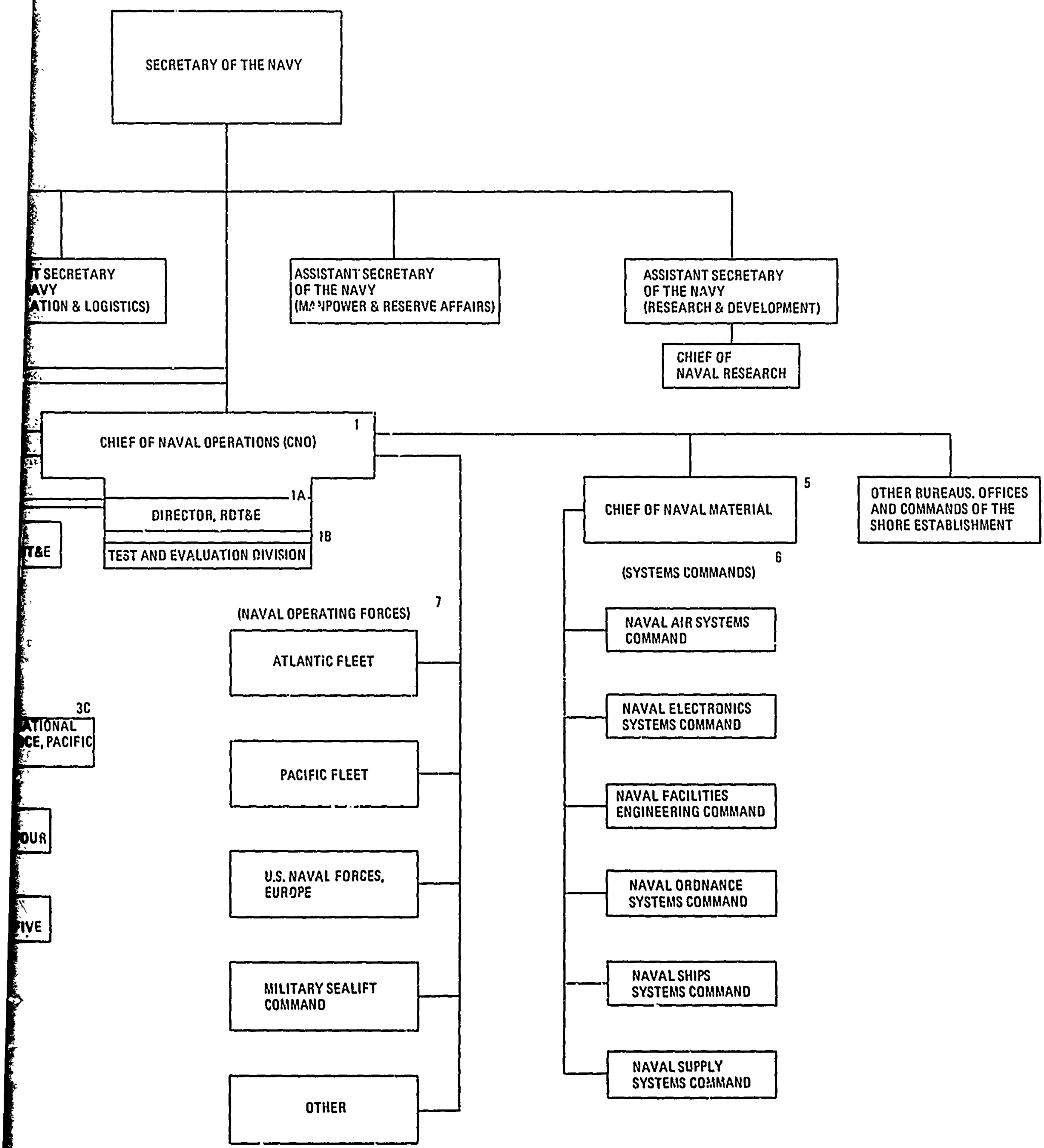


Figure D-1. Navy Organization for OT&E

Block Number	OT&E Functional Descriptor
3 (cont)	Provides results of OT&E to Defense Systems Acquisition Review Council (DSARC) and to other designated levels of review as directed by the CNO. Conducts operational tests on systems, including ships and aircraft. Evaluates operational effectiveness, suitability, and capability of tested systems. Reports results to the CNO. Assists development agencies (normally the System Commands) in development test and evaluation (DT&E) as appropriate. Reports results of assists and independent assessment to both the development agency and the CNO. Review management acquisition plans for new systems and reports to the CNO on the adequacy of the test plans in addressing critical questions and issues. Monitor and report on other T&E efforts (e.g., OT&E conducted by and within a command) as directed by the CNO.
3A	(VX-1) Air Test and Evaluation Squadron ONE: (Atlantic) Primary unit for the OT&E of Air antisubmarine warfare systems and associated tactics.

Block Number	OT&E Functional Descriptor
3B	New London Test and Evaluation Detachment: Prosecutes OT&E projects as assigned. Assists in scheduling fleet services in support of DT&E. Provides liaison with the development community in the general vicinity of New London, Connecticut.
3C	Deputy Commander Operational Test and Evaluation Force, Pacific: Structured in parallel with COMOPTEVFOR. Plans, coordinates and prosecutes OT&E with Pacific Fleet units.
3D	(VX-4) Air Test and Evaluation Squadron FOUR: (Pacific) Primary unit for OT&E of fighter aircraft, air-to-air weapon systems, and associated tactics.
3E	(VX-5) Air Test and Evaluation Squadron FIVE: (Pacific) Primary unit for OT&E of attack aircraft, air-to-ground weapon systems, and associated tactics.

Block Number	OT&E Functional Descriptor
4	Development Center, Marine Corps Development and Education Command:
	Serves as counterpart organization of COMOPTEVFOR as the independent OT&E test agency within the Marine Corps.
5	Chief of Naval Material:
	Designates the Project Manager for major defense systems, or otherwise assigns T&E responsibilities for lesser acquisition programs in order to:
	Coordinate and execute a T&E program responsive to operational questions and issues.
	Prepare a Test and Evaluation Master Plan (TEMP) early in the acquisition process to coordinate and integrate test scheduling and test accomplishment.
	Provide T&E technical support.
	Determine requirements for assistance of operating forces (which involves COMOPTEVFOR) in T&E and apprise the CNO.
6	Developing Agency (normally the appropriate Systems Command):
	Coordinate with COMOPTEVFOR on T&E planning and support

Block Number	OT&E Functional Descriptor
6 (cont)	Furnish material and technical support, including spare parts and special test equipment. Provide for installation and removal of system or equipment to be tested. Provide necessary funding to activities supporting DT&E.
7	Operating Forces Provide T&E support as requested by the CNO or the Commandant of the Marine Corps, as appropriate.

Identification of Major Participants in System Acquisition. Within the Navy, the major participants in the acquisition process for systems and equipment and the responsibilities assigned are presented as follows:

		Responsibilities:
a.	The Assistant Secretary of the Navy, (Research and Development) (ASN(R&D))	All matters related to research, development, engineering, test and evaluation within the Navy. Manages the Navy RDT&E appropriation.
b.	Chief of Naval Operations Deputy Chiefs of Naval Operations	Determine requirements State major characteristics Establish programs Act as program sponsors for procurement, modernization and alteration of systems.

	Responsibilities (cont)
	Appoint "program coordinator"
c. Director, Research, Development, Test and Evaluation (DRDT&E)	Focal point in Office of CNO for test and evaluation.
	Assists ASN (R&D) in coordination and integration of Navy RDT&E program.
	Acts as appropriation sponsor for Navy RDT&E.
	Manages the planning and reporting procedures for the conduct of the RDT&E program.
	Coordinates, formalizes require- received from DCNO's.
	Appraises the progress of RDT&E effort.
	Recommends projects for curtailment, suspension, or cancellation in favor of more worthy acquisition programs.
	Exercises cognizance over planning, conduct and reporting of all DT&E and OT&E.
	Formulates RDT&E program objectives and the annual budget for RDT&E.

d. Chief of Naval Material (CHNAVMAT)

Translates operational requirements from CNO into hardware systems.

Manages technology base development effort.

Appraises CNO of new capabilities made possible by advances in science and technology.

Develops detailed plans for RDT&E projects in response to requirements.

Designates the Project Manager for systems acquisition.

7. SIGNIFICANT DOCUMENTS IN THE ACQUISITION PROCESS

The significant documents peculiar to the Navy acquisition process, and the normal order of introduction of such documents into a given acquisition process, is presented as follows:

a. General Operating Requirements (GORs) - A requirements document which is prepared and annually updated by the Director, Research, Development, Test and Evaluation (DRDT&E) for each warfare and support area to forecast operational capability requirements for the future and provide orientation for technological research.

b. Tentative Specific Operational Requirement (TSOR) - A requirements document which is prepared by DRDT&E to identify operational requirements and to formally request the generation of a Proposed Technical Approach (PTA) by CHNAVMAT. The TSOR neither establishes firm requirements nor authorizes development programs; however, it does initiate the exploratory analysis for the technical solution to an operational need.

c. Proposed Technical Approach (PTA) - A document provided to CNO by CHNAVMAT in response to a TSOR. The PTA provides the alternative technical solutions to the operational requirement stated in the SOR and contains inputs provided by appropriate Systems Commands or responsible Bureaus or offices. PTA's may also be submitted as unsolicited proposals in response to GORs.

d. Specific Operational Requirements (SORs) - A requirements document prepared by DRDT&E to formally state the need for development of new or improved capabilities and to authorize the conduct of an Engineering Development or Operational Systems Development Project. The SOR is addressed to CHNAVMAT and requests the generation of a Technical Development Plan (TDP).

e. Technical Development Plan (TDP) - A planning document prepared by CHNAVMAT, with inputs from Systems Commands, responsible Bureaus and Offices in response to a SOR. The TDP documents the actions, procedures and resources required in attaining the operational capability described in the SOR. For programs falling below the DCP thresholds of DOD Directive 5000.1, the TDP serves to plan and control such programs within the Navy. (The SOR and the TDP provide fundamental documentation for the evaluation of the system by OPTEVFOR.)

f. Test and Evaluation Master Plan (TEMP) - A planning document prepared by the Project Manager and/or the Ship Acquisition Program Manager under the direction of CHNAVMAT which includes comprehensive time-phased requirements for subsystem and system testing. The TEMP specifies the demonstrations required for the achievement of program objectives to support decision milestones and should also address the requirements for FOT&E. OPTEVFOR is responsible for the review of the TEMP for adequacy of proposed testing and the relevance of proposed

testing and objectives to critical questions and issues and for submitting review comments and recommendations to the CNO. From the TEMP, OPTEVFOR develops both the detailed project plan for IOT&E and recommendations for the implementation of Follow-on OT&E.

g. Project Master Plan (PMP) - The PMP defines a management approach for acquiring a system or equipment. It is a compilation of planning documents which places in context the plans, schedules, costs and scope of work and resources to be provided by each participating organization. Since planning through full-scale development is covered in the TDP, the PMP extends the project objectives by emphasizing production, deployment, and support. The Project Manager is responsible for the preparation of the Master Plan.

8. IMPACT OF DOD DIRECTIVES 5000.1 AND 5000.3

During the period 1970-1973, sweeping changes were made by the Secretary of Defense (SECDEF) in the approach to the acquisition of new systems and equipment within the Department of Defense (DOD) and in the organization for and conduct of test and evaluation. These changes were begun in 1970, initiated by directive memoranda, and were later formalized in two highly significant DOD Directives, identified as DOD Directive 5000.1 (issued in 1971); subject: "Acquisition of Major Systems", and DOD Directive 5000.3 (issued in 1973); subject: "Test and Evaluation."

These directives impacted the approach to system acquisition and test and evaluation within each of the military services. An essential requirement set forth in DOD Directive 5000.3 dealt with the establishing of an independent test agency within each military component of the DOD. Within the Navy, operational test and evaluation is and has been for many years, conducted by a field agency (COMOPTEVFOR) independent of the developing command. However, in other areas, DOD Directives 5000.1 and 5000.3 required changes in Navy procedure and policy, the more significant of which were:

a. The realignment of the acquisition phases recognized by the Navy to conform to the acquisition phases set forth in DOD Directive 5000.1.

b. Changes in acquisition program management and documentation in order to satisfy DSARC review requirements and the key decision milestones of SEC DEF review and decision.

c. The application of specific management principles, set forth in DOD Directives 5000.1 and 5000.3, to all acquisition programs.

d. The resolving of major risk areas as early as practicable in ship acquisition programs, changed as follows:

(1) In the acquisition of a new ship class involving a new hull or propulsion - type, the hull and propulsion are to be prototyped and tested before following ships are started. In addition, the integrated combat systems normally are to undergo successful DT&E and an initial phase of IOT&E on land based test sites or other ships before following ships are approved.

(2) The lead ship trials (i.e., tests conducted on the first ship of a new class) are to be used by both the developer and COMOPTEVFOR to get the earliest overall evaluation of the ship.

Navy Implementation of DOD Policy For Acquisition and Test. The Navy directives listed below implement the DOD policies set forth in DOD D 5000.1 and 5000.3, and within the Department of the Navy, establish policy, provide guidance and assign responsibility concerning system/equipment acquisition and test and evaluation:

a. SECNAV INSTRUCTION 5000.1; subject: "System Acquisition in the Department of the Navy."

b. OPNAV INSTRUCTION 3930.8; subject: "Assignment and prosecution of test and evaluation projects."

c. OPNAV INSTRUCTION 3960.8; subject: "Test and evaluation of Navy systems and equipments."

d. OPNAV INSTRUCTION 5440.47; subject: "Mission and functions of Operational Test and Evaluation Force (OPTEVFOR)

(Note: The acronyms "SECNAV" and "OPNAV" stand for Secretary of the Navy and Office of the Chief of Naval Operations, respectively).

DOD Directives 5000.1 and 5000.3 also resulted in an adjustment of events in the Navy schedule of test and evaluation to align such events with the DSARC decision milestones. Figure D-2, "DSARC and Test and Evaluation Schedules" and Figure D-3, "Naval Aircraft and Missile Acquisition T&E Cycle" show the current relationship of program development, including DT&E, IOT&E and FOT&E, to the phases of the acquisition cycle and the DSARC events.

9. PRINCIPAL NAVY TEST AND EVALUATION (T&E) FACILITIES

Pacific Missile Range, Point Mugu, California

Atlantic Fleet Weapons Range, Roosevelt Roads, Puerto Rico

Atlantic Underseas T&E Center, Andros Island, Bahamas

Naval Weapons Center, China Lake, California

Naval Air Propulsion Test Center, Trenton, New Jersey

Naval Air Test Center, Patuxent River, Maryland

Naval Aerospace Recovery Facility, El Centro, California

Naval Air Test Facility, Lakehurst, New Jersey

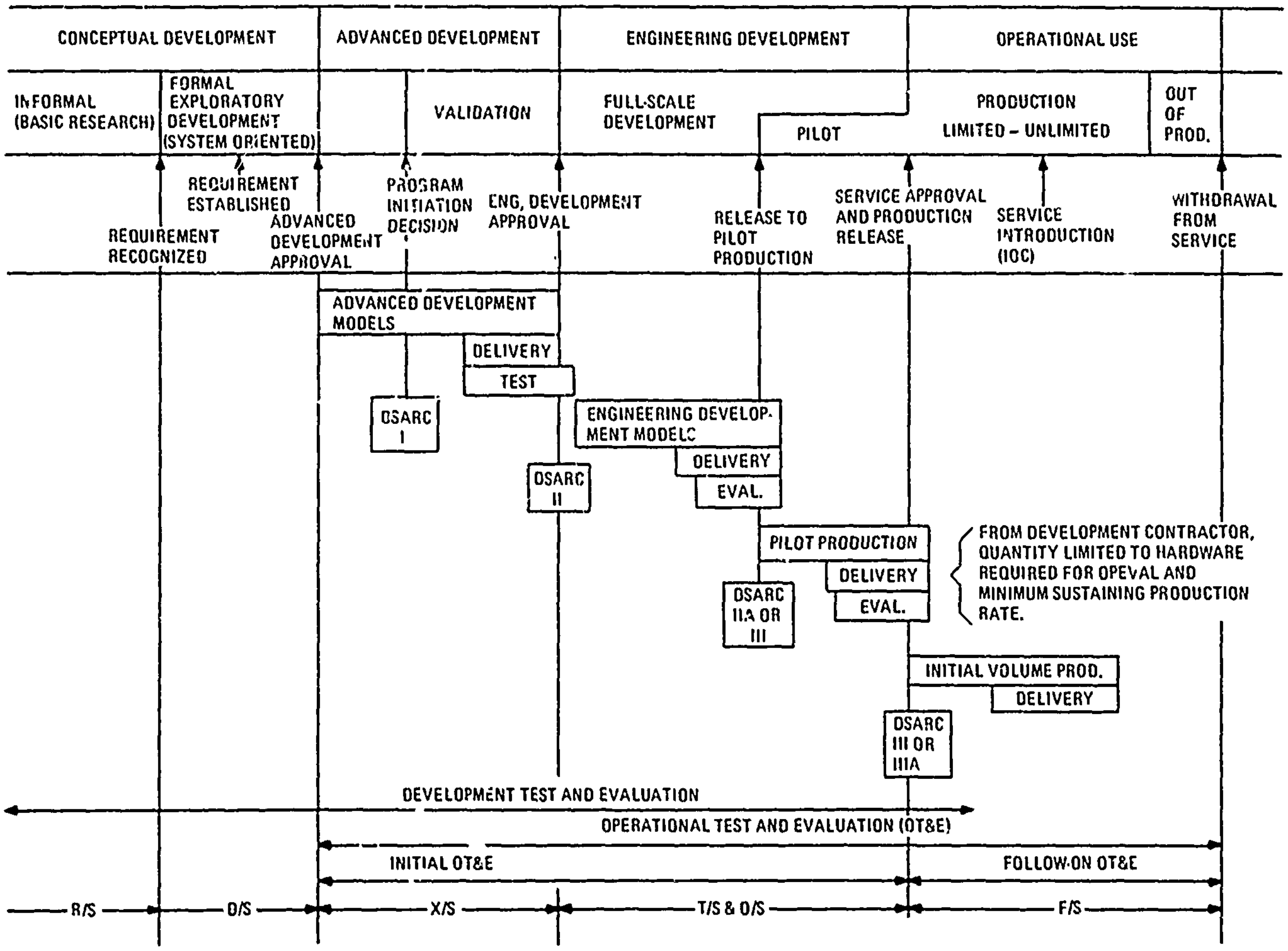


Figure D-2. DSARC and test and evaluation schedules.

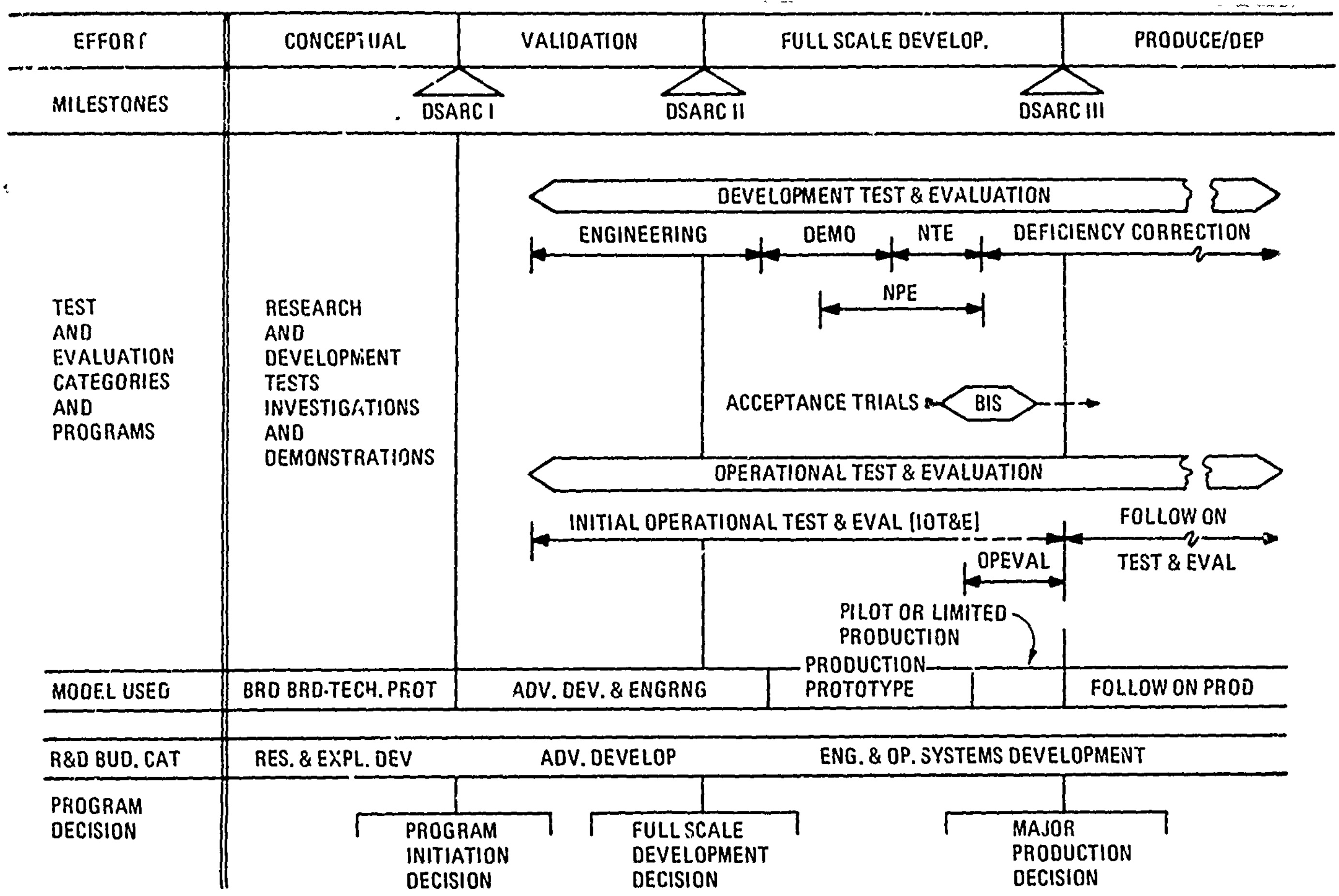


Figure D-3. Naval aircraft and missile acquisition T&E cycle.

Appendix E

THE COMPARISON OF OT&E BETWEEN THE ARMY, NAVY AND AIR FORCE

General. The overviews of Operational Test and Evaluation in the Army and the Navy appear herein as Appendices C and D, respectively. The main text of this document (Chapters 1 through 15) offer a broad scope view of like activities within the Air Force. For the Air Force reader, the ability to compare the OT&E activity between services affords an opportunity to identify similarities and contrasts and as is true in all technical activity, provides a healthy cross-fertilization of organizational and procedural concepts and philosophy.

Joint Tests. From time to time, the Office of the Secretary of Defense (OSD) prescribes a Joint Test involving the Air Force and one or more of the other military components within the Department of Defense (DOD). The organizations of the Army and Navy for conducting OT&E and the related information on each Service's activity should prove useful to the Air Force personnel involved in the coordination of Joint Test activity who perform the interface function with other Services.

Summary View. In addition to the detailed information on Army and Navy OT&E provided in Appendices, Figure E-1. "ARMY/NAVY/AIR FORCE COMPARISON" provides a summary view and comparison between the Services.

	US Air Force	US Army	US Navy
1 Designation of major field agency responsible for OT&E	1 Air Force Test and Evaluation Center (AFTEC) HQ AF Test & Evaluation Center Kirtland AFB NM 87115 Autovon - 829-4118 Commercial - (505) 264-4118	1 US Army Operational Test and Evaluation Agency (OTEA) CG, HQ US Army Operational Test and Evaluation Agency Fort Belvoir VA. 22060 Autovon - 354-4735 Commercial (703) 664-4735	1 Operational Test and Evaluation Force (OPTEVFOR) Commander, Operational Test and Evaluation Force Norfolk VA. 23511 Autovon - 690-5061 Commercial - (804) 444-5187
2 Manning levels (authorized)	2 Military - 166 Civilian - 42	2 Military - 148 Civilian - 102	2 Military - 1329 Civilian - 46
3 Reporting chain:	3 Reports direct to the CSAF Interfaces with AF/X00W to provide OT&E staff surveillance.	3 Reports direct to CSA.	3 Reports to CNO on OT&E matters (a) Logistic support provided by CINCLANTFLT/CINCPACFLT/ CINCUSNAVEUR as appropriate. (b) Admin support provided by CINCLANTFLT.
4 Principal directives/regulations concerning systems acquisition and OT&E.	4 • AFR 800-2 "Program Management" • AFSCP 800-3 "A Guide For Program Management" • AFLCM 800-1 "Program Management" • AFR 80-14 "Test and Evaluation" • AFR 23-36 "Air Force Test and Evaluation Center (AFTEC)"	4 • AR 1000-1 "Basic Policies For Systems Acquisition By the Department of the Army" • AR 70-10 "Test and Evaluation During Development and Acquisition of Material" • AR 10-4 "US Army Operational Test and Evaluation Agency" • DA Letter of Instructions (LOI) for Management of Joint User Testing Programs dated 2/28/74	4 • SECNAVINST 5000.1 "System Acquisition in the Department of the Navy" • OPNAVINST 3930.8 "Assignment and Prosecution of Test and Evaluation Projects" • OPNAVINST 3960.8 "Test and Evaluation of Navy Systems and Equipments" • OPNAVINST 5440.47 "Mission and Functions of Operational Test and Evaluation Force.
5 Field activities directly under OT&E Test Agency.	5 All facilities support and some personnel support provided by designated MAJCOMS.	5 None. OTEA contributes deputy test director and ~3-5 key personnel to test directorate at the command conducting test.	5 • Three (3) Air, Test and Evaluation Sqdns (VX1, 4, and 5) • Deputy COMOPTEVFOR PAC. San Diego • New London Test and Evaluation Detachment
6 Comparison of characteristics of OT&E	6 OT&E (a) Development of optimum tactics, techniques, procedures and concepts. (b) Evaluation of reliability, maintainability and operational effectiveness and suitability. (c) Testing under realistic operational conditions. IOT&E (a) Accomplished prior to first major production decision. (b) Determine operational effectiveness and suitability. FOT&E (a) Subsequent to receipt of production items. (b) "IOT&E will not obviate the need for FOT&E" (c) "Focuses on the operational and employment aspects of the system". AFR 80-14	6 OT (a) Representative user troops (b) Realistic operational environment (c) Provide data to estimate military utility, operational effectiveness, and operational suitability (including compatibility, interoperability, reliability, availability, maintainability, integrated logistics support and training requirements. (d) (From users viewpoint): (1) System desirability considering systems already available. (2) Operational benefits/ burdens. (3) The need for modifications to a system. (4) The adequacy of doctrine, organization, operating techniques, tactics and training for employment/ logistic support. AR 10-4	6 OT&E (a) Participated in or performed by operational personnel. (b) Focuses on operational effectiveness and suitability (including reliability, compatibility, interoperability, maintainability and supportability. (c) Development of optimum operational tactics. IOT&E (a) Accomplished prior to DSARC III or comparable CNO or CHNAVMAT major production decision point. (b) Assessment of operational effectiveness and suitability. FOT&E (a) Continuing OT&E conducted in operational environment by operational personnel using production systems. (b) Verify system performance. (c) Validating correction of deficiencies previously identified. (d) Refining tactical employment doctrine and requirements for personnel and training.

of OT&E	(a) Development of optimum tactics, techniques, procedures and concepts. (b) Evaluation of reliability, maintainability and operational effectiveness and suitability. (c) Testing under realistic operational conditions. IOT&E (a) Accomplished prior to first major production decision. (b) Determine operational effectiveness and suitability. FOT&E (a) Subsequent to receipt of production items. (b) "IOT&E will not obviate the need for FOT&E" (c) "Focuses on the operational and employment aspects of the system". AFR 80-14	(a) Representative user troops. (b) Realistic operational environment (c) Provide data to estimate military utility, operational effectiveness, and operational suitability (including compatibility, interoperability, reliability, availability, maintainability, integrated logistics support and training requirements. (d) (From users viewpoint): (1) System desirability considering systems already available. (2) Operational benefits/burdens. (3) The need for modifications to a system. (4) The adequacy of doctrine, organization, operating techniques, tactics and training for employment/logsitic support. AR 10-4 IOT&E OT I: Examine hardware or components to provide an indication of utility and worth to the user. In case of "off the shelf" competitors, execute a "fly-off". OT II (a) Examination of engineering development/preproduction prototype equipment prior to initial production decision. (b) Assess military utility, operational effectiveness and operational suitability including RAM (reliability, availability, maintainability) in a realistic environment. OT III (a) Examine initial production items. (b) Assurance of operational suitability (c) All operational critical questions/issues resolved. Benefits/burdens identified. AR 70-10 FOT&E: (See footnote ‡)	(a) Participated in or performed by operational personnel. (b) Focuses on operational effectiveness and suitability (including reliability, compatibility, interoperability, maintainability and supportability. (c) Development of optimum operational tactics. IOT&E (a) Accomplished prior to DSARC III or comparable CNO or CHNAVMAT major production decision point. (b) Assessment of operational effectiveness and suitability. FOT&E (a) Continuing OT&E conducted in operational environment by operational personnel using production systems. (b) Verify system performance. (c) Validating correction of deficiencies previously identified. (d) Refining tactical employment doctrine and requirements for personnel and training. (e) "May be initiated using pilot or pre-production systems .... until the production units are available." OPNAVINST: 3960.8
7 Authority for assignment of OT&E projects.	7 (a) CSAF (AF/X60) (b) AFTEC may recommend to CSAF those OT&E projects which AFTEC or MAJCOMs should conduct.	7 (a) DCSRDA determines "major systems". (b) Tasking of other commands for "User testing" (which includes OT&E) is directed by TSARC.	7 (a) CNO (Director, RDT&E (OP-098)) (b) CNO may authorize COMOPTEVFOR to further assign projects to other commands when advantages are apparent.
8 Identification of master test document for scheduling multi-program OT&E.	8 Air Force OT&E Master Program	8 US Army Five Year Test Plan (FYTP)	8 None.
9 Master test document for specific (major or selected nonmajor) IOT&E program.	9 Section 5 (Test and Evaluation) of Program Management Plan (PMP).	9 Coordinated Test Program (CTP) (Section IV of the project manager's Development Plan (DP)).	9 Test and Evaluation Master Plan (TEMP)

Figure E-1. ARMY/NAVY/AIRFORCE comparison (Sheet 1 of 2)

	evaluation) (l) Final Test Report (m) AFTEC Commander's Summary (final independent operational evaluation by AFTEC)	evaluation by OTEA) (l) OT III/DT III Test Reports (independent operational evaluation by OTEA)	
12 Commands conducting OT.	12 (a) AFTEC – (all major and selected nonmajor systems) (b) Designated AF MAJCOMS – (non major systems)	12 (a) OTEA – (all major and selected nonmajor systems) (b) Combat Development Experimentation Command (CDEC) – (nonmajor systems) (c) Modern Army Selected Systems Test Evaluation and Review (MASSTER) – (non major systems) (d) Other Major Commands – (non major systems) (e) Test boards (nonmajor systems)	12 COMOPTEVFOR
13 Commands performing evaluation of OT&E directed by Service HQ.	13 The command conducting OT performs the evaluation subject to review, approval, and further evaluation by AFTEC.	13 OTEA performs the evaluation.	13 COMOPTEVFOR performs the evaluation.
14 Use of operations analysts within test agency	14 (a) Select technically feasible and analytically sound methods for OT&E. (b) Prepare inputs to test directive. (c) Provides operations analyst point of contact for each AFTEC test. (d) Analyze data derived from tests. (e) Periodically check to see that data collection procedures are producing the desired data. (f) Work with data process personnel to ensure that data processing techniques are compatible with test objectives. (g) Prepare the technical analysis portions of the report of OT&E. (h) Assist in planning the DSARC briefing.	14 (a) Assist in development of initial test designs to include: (1) Listing of test measures and conditions. (2) Integration of operational testing requirements into development testing. (3) Structuring of separate operational testing. (b) Assist in the development of Outline Test Plans (OTPs) (c) Aid in the development of Test Plans. (d) Provides schedule for detailed analysis planning and data collection plan. (e) Analyze data results of testing. (f) Prepare the technical analysis portion of the test report and analysis inputs to the OTEA evaluation and Congressional Reports. (g) Assist in planning DSARC briefings.	14 (a) Assist Project Officer by the assignment of a project analyst and material analyst to each OT&E project. Analysts serve to advise and assist in the following areas: (1) Applicability and adequacy of performance criteria. (2) Development of MOE's (Measures of Effectiveness) (3) Test Design (4) Data collection (5) Data reduction (6) Data analysis (b) Prepare portions of test design, test plan and test report/evaluation concerning (1) through (6) immediately above. (c) Review and determine applicability of data from other sources.

‡NOTE: FOT&E is barely mentioned in Army Directives concerning OT&E policy. The evidence indicates that FOT&E is rarely undertaken. New systems are preferred over modifications to old systems and where a significant modification to an old system is accepted, the modification is developed and tested with essentially the same management procedures as are prescribed for a new system. At the time of this study, this philosophy was under consideration for change.

Figure E-1. ARMY/NAVY/AIRFORCE comparison (Sheet 2 of 2)

		US Air Force	US Army	US Navy
10	Views on combined DT&E/ IOT&E (major & selected nonmajor systems)	10 (a) Operating/supporting commands participation is mandatory during development testing, beginning with test planning prior to Program Decision. (b) Operational inputs will be utilized in (test) planning documents developed by the Implementing Command. (c) To maximum degree possible: (1) IOT&E will be accomplished by operating personnel of the appropriate Operating/ Supporting Commands. Realistic operational environment. Use pilot or early production items. (Separate IOT&E) (d) Where adequate test data can be secured from combined DT&E/IOT&E, combined tests may be employed. Added considerations for combined testing are cost, time. (ARF 80-14) (e) Separate evaluations required from Operating/Supporting Commands to assist in Production Decision.	10 (a) "The OT&E process will be independent and should normally be separate from the DT&E process". (AR 70-10) (b) "May be combined where separation causes delay involving unacceptable risk or unacceptable acquisition costs". (AR 70-10) (c) Testing is usually phased as DTI, II, and III and OT I, II, and III. Usually OT I is run in combination with DT I. Army attempts to keep OT II and III and DT II and DT III separate. Sometimes combined; however. (d) Each phase of OT testing results in a separate evaluation by OTEA, timed to provide evaluation to the decision body (DSARC, ASARC or IPR) at decision milestones. (e) DT II, OT II, and DT III are conducted on all development items/systems as a minimum. Non-development items normally undergo only DT III.	10 (a) OPTEVFOR participates in the TECH EVAL (DT&E) which is planned, conducted by the Developing Agency with OPTEVFOR inputs. OPTEVFOR provides independent operational assessment to CNO. (b) OPTEVFOR responsible for planning, conducting, reporting OPEVAL, conducted after TECHEVAL and prior to Production Decision milestone. TECH EVAL + certain OPEVAL data form basis for independent evaluation by OPTEVFOR.
11	Document chain in system acquisition (related to OT&E)	11 (a) ROC (formally states a Required Operational Capability) (b) Action Directives (early PMD). (c) Program Advocacy Documents (studies, analyses, etc.) (d) DCP - (OT&E inputs provided by AFTEC) (e) PMD - Initial Test Directive - (OT&E inputs provided by AFTEC) (f) Test objectives annex to the PMD - (OT&E inputs provided by AFTEC) (g) HQ USAF Test Directive (Drafted by AFTEC; based on c, d, e, f) included within the Air Force OT&E Master Program (h) AFTEC Commander's Estimate (i) AFTEC Test Plan (j) Program Management Plan (contains DT&E, OT&E test plans) (k) Interim Test Reports (interim evaluation) (l) Final Test Report (m) AFTEC Commander's Summary (final independent operational evaluation by AFTEC)	11 (a) ROC (formally states a Required Operational Capability) (b) Charter for special task force (c) OTEA evaluation plan. (d) Program Advocacy Documents (Concept Formulation Package) (e) DCP (OT&E inputs provided by OTEA) (f) Final Report of Special Task Force (contains plan for OT&E provided by OTEA). (g) OTEA or TRADOC Outline Test Plans. (h) Test Plans, included within the Army Five Year Test Plan (FYTP) (i) Development Plan (contains "Coordinated Test Program" of all DT&E and OT&E testing of system) (j) OT I/DT I Test Reports (independent operational evaluation by OTEA) (k) OT II/DT II Test Reports (independent operational evaluation by OTEA) (l) OT III/DT III Test Reports (independent operational evaluation by OTEA)	11 (a) SOR (formally states a Specific Operational Requirement) (b) Technical Development Plan (TDP) - (program advocacy documents including OT&E inputs by OPTEVFOR.) (c) OPTEVFOR evaluation plan. (d) DCP - (OT&E inputs provided by OPTEVFOR) (e) Test and Evaluation Master Plan (TEMP) - (contains DT&E and OT&E plans for a specific program.) (f) Project Master Plan (PMP) - (contains TEMP) (g) TECHEVAL Plan - (the developing agency's plan for DT&E) (h) Tech Eval Report (i) OPTEVFOR's independent assessment of the Tech Eval. (j) OPEVAL Plan - (the OPTEVFOR plan for OT&E) (k) OPEVAL Report - (independent operational evaluation by OPTEVFOR)
12	Commands conducting OT.	12 (a) AFTEC - (all major and selected nonmajor systems) (b) Designated AF MAJCOMS - (non major systems)	12 (a) OTEA - (all major and selected nonmajor systems) (b) Combat Development Experimentation Command (CDEC) - (nonmajor,	12 COMOPTEVFOR

Appendix F

DEVELOPMENT OF STATISTICAL DESIGN

This Appendix discusses the principal areas of consideration which go into the development of the statistical design for a test. The statistical design for the test (1) dictates questions which may be addressed in the test program that are based upon an objective analysis of the experimental data, (2) suggests the resources which are necessary for the test, (3) defines the experimental conditions that are appropriate for each test trial, and (4) defines the data analysis plan for processing the experimental data after it is collected, verified and reduced.

1. IDENTIFY DEPENDENT VARIABLE CHARACTERISTICS

The observations to be made in the test are the dependent variables in the statistical design. The form of response that each dependent variable assumes must be identified to facilitate an appropriate statistical design. To illustrate, the response could be either a continuous variable over a range of values, (e.g., miss-distance of a missile from a target) or the response could be one of two values, (e.g., success or failure, yes or no, go or no-go), or the response could take on a series of discrete values, (e.g., judges' rating, using the integers 1, 2, ..., 10, of an aircrew's adequacy of or performance for specific functions).

Knowledge of the type of probability distribution to expect for dependent variables, (e.g., Normally distributed, uniformly distributed) should be identified to provide a basis upon which to structure a statistical design. For example, some statistical designs assume a Normally-distributed variable. When there is lack of adequate knowledge about the probability distribution for a variable, non-parametric techniques may be required in the statistical design.

2. IDENTIFY LEVELS AND COMBINATIONS OF TEST FACTORS (INDEPENDENT VARIABLES)

The test factors that have been identified previously in the test planning process are the independent variables in the test and must be analyzed to further define the levels

(or range of values) they may assume in the test. The various combinations of test factors that are present in the test must be considered also since this will influence the scope of the experiment.

One should be aware as to whether a test factor is qualitative, (e.g., an aircrew or aircraft) or quantitative, (e.g., the altitude of weapon release). For a particular test, a test factor may be held fixed at a series of specified levels, (e.g., weapon release altitude at 10,000, 20,000, 30,000 and 40,000 feet) or permitted to vary at random within a range of values. When the principal set of combinations of test factors that are to be examined in the test have been identified, the test planner has one of the basic ingredients with which to develop a statistical design.

3. IDENTIFY TYPE OF STATISTICAL PROBLEM

The test planner must translate the basic test objectives and questions into specific types of statistical problems which can be addressed with the experimental data that is gathered during the test trials. "Statistics" is an activity that assists in the decision-making process under the conditions of uncertainty or random variation in the outcome of a particular observation. The uncertainty or random variation does not mean that there is complete ignorance of the system under test; simply, that since random variation is present in all groups of observations, the real variation due to the test factors must be detected in its presence.

There are three basic types of statistical problems into which the test objectives and questions should fall: (1) estimations, (2) comparisons, and (3) determining relationships between variables. An example of an estimation problem would be to estimate the average miss-distance from a target by a missile and the miss-distance interval which would have a 95 percent chance of including the true population mean miss-distance. An example of a comparison would be to determine if system A is better than system B. An example of determining a relationship would be the derivation of an equation which described the average target detection time for an aircrew as a function of the target size, movement, and color contrast. There is an element of

overlap in the mechanics of the mathematics between the determination of a relationship and the estimation and comparisons problems; however, it is convenient to categorize them separately since the orientation to the test designers' problems is different.

4. MATCH PROBLEM TYPE TO CANDIDATE SOLUTION TECHNIQUES

Having identified the general type of statistical problem for the test, the next step is to derive candidate statistical designs for the test that will solve the statistical problem. There are two general categories of statistical design techniques. The first category, called parametric techniques, is the most commonly used and relates to the situation where characteristics of the probability distribution for the dependent variable are known or assumed, (e.g., Normally distributed). The second category, called non-parametric techniques, relates to the situation where characteristics of the probability distribution for the dependent variables are unknown or ignored.

Parametric Techniques

The process of inferring something about a population on the basis of a sample drawn from the population is called statistical inference. Most statistical theory is based on the assumption that the samples drawn are random samples. Each member of the population has a known probability of being included in the sample and that the pattern of variation in the population is not changed by extracting the sample. Parametric statistics entails the characterization of the population under study by one or more parameters, (e.g., mean and variance).

Estimation of Mean, Variance, and Confidence Interval. Sample statistics; or simply statistics, are computed from the test data observations to estimate the parameters which characterize the population from which they are drawn. The population parameters are denoted by Greek letters and include the mean (μ), the variance, (σ^2) and the standard deviation (σ). The sample mean ($\bar{x}$) is computed by Equation (1)

$$\bar{x} = \sum_{i=1}^{n} x_i \Big/ n \quad , \tag{1}$$

where there are n observations of x in the sample. The sample variance (s^2) is given by Equation (2)

$$s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \Big/ (n-1). \tag{2}$$

The denominator in Equation (2) is n-1 rather than n to provide an unbiased statistic. An unbiased statistic is one whose expected or average value taken over an infinite number of similar samples will be the true value of the population parameter being estimated. For example, $\bar{x}$ and s^2 are unbiased point estimates of μ and σ^2, respectively.

An <u>interval estimate</u> of a population parameter consists of the interval between two values of a function of the sample statistic that is asserted to include the parameter under discussion. The range of values between these two extreme values is the <u>confidence interval</u> for the parameter, since its width may be determined from the degree of confidence that is assumed when one asserts that the parameter lies within the interval. The <u>confidence limits</u> are the end points on the confidence interval and are determined from the observed sample statistics, the sample size, and the degree of confidence desired in the application. For example, if one assumes a Normally distributed variable, a 95 percent confidence interval for μ (the population mean) is given by Equation (3)

$$95\% \text{ confidence interval for } \mu = \bar{x} \pm 1.96\, \sigma/\sqrt{n}\ , \tag{3}$$

where 1.96 is taken from the standard Normal distribution tables and corresponds to the distance (in terms of the number of standard deviations), in both directions, from the mean for the familiar bell-shaped curve that would include 95 percent of all observations. The population standard deviation (σ) is assumed known when using Equation (3). When σ is unknown, the "Student" t distribution is used and, hence, the confidence interval is given by Equation (4)

$$100(1-\alpha)\ \% \text{ confidence interval for } \mu = \bar{x} \pm t_{(1-\alpha/2)}\, s/\sqrt{n}\ , \tag{4}$$

where t is taken from tables for the Student t distribution, entering with parameters n-1 (degrees of freedom) and $1-\alpha/2$ (confidence level). The following expression permits the computation of the number of degrees of freedom:

$$\text{Degrees of freedom} = \begin{pmatrix}\text{number of}\\ \text{independent}\\ \text{observations}\end{pmatrix} - \begin{pmatrix}\text{number of inde-}\\ \text{pendent parameters}\\ \text{estimated in com-}\\ \text{puting variation}\end{pmatrix}$$

In this example the number of independent observations is n and an estimate of the population mean was used in computing the variation by equation (2)

Comparing Two Means. The discussion for the comparison of two means illustrates many of the basic principles that apply to more complex statistical problems. A statistical hypothesis is an assumption about the population being sampled. When comparing the means of two populations, it can be hypothesized that the difference between the means is zero. This is expressed as H_0: $\mu_1 - \mu_2 = 0$.

A test of hypothesis is a rule by which a hypothesis is either accepted or rejected. The rule is usually based on sample statistics, called test statistics when they are used to test hypotheses. The critical region of a test statistic consists of all values of the test statistic where the decision is made to reject H_0.

A decision made to either accept or reject a hypothesis is subject to two kinds of possible errors since the test is based on sample statistics derived from n observations. If a hypothesis is really true although it is rejected on the basis of the test statistics, a Type I error has been committed. The Greek letter α is used to denote the probability of having a Type I error occur. In our example when H_0 is rejected, an alternative hypothesis (that the means are not equal) is accepted and this hypothesis is denoted by H_1: $\mu_1 \neq \mu_2$.

If a hypothesis (H_0) is accepted when it is not really true, a Type II error has been committed. The probability of having a Type II error occur is denoted by β. The α and β error probabilities are the risks of making incorrect decisions based on the test statistics. A principal objective in statistical test design is to make α and β as

small as possible while balancing the corresponding expenditure of test resources to increase the sample size n (increasing n tends to reduce the β error for a fixed level of α).

In typical test design applications, the value of α is set at a particular value, (e.g., $\alpha = 0.05$ or $\alpha = 0.10$). For the comparison of two means, the test designer should determine the probability (β) of not detecting a specified difference between the population means when, in fact, that difference is the true difference. For example, suppose that the requirement exists to detect a difference in means if there is a true difference of 25 percent with a probability of 0.90 ($\beta = 0.10$, the probability of not detecting the difference). Thus, knowing α, β for a specific difference in true means, and an estimate for σ or s, the analyst may consult statistical tables to determine the proper sample size n for the test.

There are either of two test statistics to employ in the comparison of two population means. If the population variance is known or the sample variance is estimated from a sufficiently large sample, the standard normal variable (Z) may be used:

$$Z = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{5}$$

The critical region corresponding to the 95 percent <u>confidence level</u> (α=0.05) is $|Z| > 1.96$. Thus, one could compute Z from Equation (5) and accept H_0: $\mu_1 = \mu_2$ if $Z < 1.96$; otherwise, accept H_1: $\mu_1 \neq \mu_2$ when $|Z| > 1.96$.

When the sample sizes are small and the population variance is unknown, the Student t-distribution is used for the test statistics

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) s^2}} \text{ where } s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} . \tag{6}$$

The critical region to test the hypothesis H_o: $\mu_1 = \mu_2$ corresponds to those values of t which are greater than the absolute table value (t_{table}) for the Student t-distribution with a specified α-error and $n_1 + n_2 - 2$ degrees of freedom. Thus, accept H_o: $\mu_1 = \mu_2$ if $|t| \leq t_{table}$; otherwise accept H_1: $\mu_1 \neq \mu_2$ if $|t| > t_{table}$.

Comparing Two Variances. Consider the situation where one desires to test whether the variances from two normal populations are equal. The hypothesis to test is H_o: $\sigma_1^2 = \sigma_2^2$. The test statistic is given by Equation (7)

$$F = \frac{s_1^2}{s_2^2}, \tag{7}$$

where s_1^2 and s_2^2 are arranged such that $s_1^2 > s_2^2$.

The critical region corresponds to those values of F which are greater than a table value for the F-distribution for a Type I error risk of α and there are n_1-1 and n_2-1 degrees of freedom in the numerator and denominator, respectively. The variance ratio or F-distribution has many applications for more complex statistical problems, (e.g., analysis of variance).

Comparing More Than Two Means. Generally, OT&E involves the comparison of more than two population means based on sample statistics. The technique usually employed to solve such problems is called analysis of variance. Fundamentally, analysis of variance (hereafter, abbreviated ANOVA) is just what the name implies - partitioning the variance of the dependent variable from an experiment into parts to test whether or not certain factors (independent variables) that were introduced into the design actually affect its value. For example, is the miss-distance of a missile system affected by the particular aircraft which releases it? Does the type of radar aboard an aircraft affect the time of target acquisition? In each case, there is interest in testing whether the factor(s) under study significantly affect the measured response variable when compared to the random variation in the process.

When the proper conditions and assumptions are present, the ANOVA technique is a powerful technique to use for statistical problems that involve the comparison of more

than two means. Basically, the efficiency of ANOVA is derived by utilizing all the observations across all combinations of test factors to estimate the experimental error or random error inherent in the process. The F-test or variance ratio is used to compare the estimated variability attributable to a test factor to the estimated experimental error and, subsequently, test for a significant effect.

Various ANOVA Models. There are numerous types of ANOVA models, each incorporates particular assumptions that describe the manner in which the test is structured and conducted. An overview of several of the principal types of ANOVA models are discussed below:

a. Single Factor. In discussing the single factor ANOVA model, many of the principles involved apply to more complex designs with only slight, but important, modifications. The single-factor experiment involves the test to see whether there is a significant difference between the levels of one factor. For example, consider the experiment where there is interest in determining whether there is a difference between four aircraft types for their effect on the radial miss-distance of an identically-launched missile at a target. Table F-1 shows the data for this example.

TABLE F-1. RADIAL MISS-DISTANCE

Aircraft Type			
I	II	III	IV
51	59	40	37
50	56	41	34
57	45	40	40
54	50	34	38
55	51	38	36

The order in which the 20 observations were taken was completely random. Therefore, the example problem has a completely randomized design. The model for the design is given by Equation (8)

$$X_{ij} = \mu + A_j + \epsilon_{ij}, \tag{8}$$

where X_{ij} is the value for the dependent variable for the i^{th} observation within the j^{th} aircraft type, A_j represents the effect for the j^{th} aircraft type, and ϵ_{ij} represents the random error that is the present in the i^{th} observation within the j^{th} aircraft type.

The model has the following assumptions:

(1) The effects are additive as shown by Equation (8).

(2) The error term ϵ_{ij} is a Normally and independently distributed random effect. It has mean value zero and its variance is the same for each level (the four aircraft types in the example) of the test factor.

(3) μ is a fixed (but unknown) parameter for the population mean.

(4) The sum of the factor level effects add to zero. This may be expressed by Equation (9)

$$\sum_{j=1}^{J} A_j = 0, \tag{9}$$

where $J = 4$ (the number of aircraft types) in our example. If the J levels of the factor are chosen at random, the A_j are assumed Normally and independently distributed with mean zero and with a common variance σ_A^2. When the A_j are set at pre-determined levels, they are called fixed and Equation (9) also applies.

The hypothesis that is tested in the single factor design is $H_0: A_j=0$ for all j. If this hypothesis is not rejected, then it is assumed that there is no effect

introduced by the type of aircraft and that each observation X_{ij} is made up of a mean μ and a random error ϵ_{ij}.

The ANOVA table for the example is shown in Table F-2. The sources of variation are the "between aircraft types effects" and the experimental error. The "Sum of squares" are computed from a set of equations that are derived from the four assumptions for the model and the observation data, X_{ij}. The "degrees of freedom" correspond to the number of observations in a sample used to estimate a parameter minus the number of parameters that are being estimated for the same sample, (e.g., there are four A_j means to estimate the average A effect.) "Mean squares is computed by dividing the sum of squares by the degrees of freedom in each row.

TABLE F-2 - SINGLE FACTOR ANOVA EXAMPLE

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	$F_{.05}$
Between aircraft types, A_j	1135.0	4-1 = 3	378.3	29.8	3.24
Experimental error, ϵ_{ij}	203.2	20-3-1 16	12.7		
Totals	1338.2	20-1 · 19			

The test statistic is the F value which is the ratio of mean squares for A_j and ϵ_{ij}. The critical region in our example is the range of F values that are larger than the table F value for $\alpha = 0.05$ and 3 degrees of freedom in the numerator and 16 degrees of freedom in the denominator. Since $F = 29.8 > 3.24$, there is at least one statistically significant difference in aircraft types. A casual look at the data in Table F-1 re-confirms this statistical decision. The determination of which combinations of aircraft types are different from each other is discussed in a later section (after ANOVA).

b. Two Factors. In an effort to further refine the experimental error (which is the yardstick by which to test for a significant effect from the levels of the test factors), a restriction may be added to the randomization in our single

factor example problem model. The "restriction" is to consider the effect that different aircrews may have on the measured variable. Table F-3 shows how the test would look if the restriction is made that every aircraft type must be used once by each of the five separate aircrews. The result of adding the restriction for aircrews is that the design is now a two factor design. Equation (10) gives the model for our example with a second factor added

$$X_{ij} = \mu + A_j + B_i + \epsilon_{ij} \tag{10}$$

where B_i represents the aircrew effect. Another way that the design may be described is to refer to the aircrews as "blocks" and that the randomization is now restricted within blocks, (e.g., each aircrew must use each aircraft type but the test order is selected in a random manner). Thus, another name for the design is "single factor randomized complete block design."

TABLE F-3 - TWO FACTOR DESIGN EXAMPLE PROBLEM

Aircrew, B_i	Aircraft Type, A_j			
	I	II	III	IV
1	X_{11}	X_{12}	X_{13}	X_{14}
2	X_{21}	X_{22}	X_{23}	X_{24}
3	X_{31}	X_{32}	X_{33}	X_{34}
4	X_{41}	X_{42}	X_{43}	X_{44}
5	X_{51}	X_{52}	X_{53}	X_{54}

The arrangement for the ANOVA table for our two factor example problem is shown in Table F-4. One can see that our original error term in the single factor design has now been broken down into two components. If there is a significant aircrew effect, then the original error term would have been quite large and there probably would have been difficulty in detecting any significant aircraft type effect.

TABLE F-4 - TWO FACTOR ANOVA EXAMPLE

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	$F_{.05}$
Between aircraft types, A_j	SS_a	4-1 = 3	$SS_a/3=MS_a$	MS_a/MS_ϵ	3.49
Between aircrews, B_1	SS_b	5-1 = 4	$SS_b/4=MS_b$	MS_b/MS_ϵ	3.26
Experimental error, ϵ_{ij}	SS_ϵ	20-7-1 = 12	$SS_\epsilon/12=MS_\epsilon$		
Totals	SS_T	19			

A comparison between our single factor design and the two factor design illustrates an important principle in the design of experiments. In the single factor design, the aircrew effect is not in the design model; hence, the aircrew effect is confounded into the experimental error. (This is acceptable when from previous experience there is strong reason to believe that the effect is negligible). In the two factor design, the experimental error is more precisely estimated, (since the potential aircrew effect is identified) but has a corresponding decrease in the number of degrees of freedom, (i.e., 16 in the single factor example versus 12 in the two factor example).

c. Latin Square. A Latin Square design is one where each level of each factor is combined only once with each level of two other factors. Consider our previous two factor design example and add a third factor (C_k, k=1, 2, 3, 4) that is the four different production lot groups from which are sent the missiles that are used in the test. Further, to illustrate the Latin Square design, the number of levels of aircrews has been reduced from five to four. Table F-5 shows the resultant 4 x 4 Latin Square design arrangement where each level of each factor occurs once and only once with each level of each of the other two factors.

TABLE F-5 - 4 x 4 LATIN SQUARE DESIGN EXAMPLE

		Aircraft Types, A_j			
		I	II	III	IV
Aircrews, B_i	1	C_1	C_2	C_3	C_4
	2	C_4	C_1	C_2	C_3
	3	C_3	C_4	C_1	C_2
	4	C_2	C_3	C_4	C_1

Missile Production Lot Groups, C_k

Some of the freedom for randomization has been lost with the Latin Square design but not all of it. For a given problem, one can select at random from tables that contain different Latin Squares design arrangements of the required size, (e.g., 4 x 4, 5 x 5).

A serious consideration that must be given to the Latin Square design is that the interaction effects of the test factors are confounded into the experimental error term. If there are interaction effects present in the test, the error term will be inflated and, thus, it will be difficult to detect other significant factor effects. Hence, a critical decision must be prior to the selection of a Latin Square design for a test as to whether there is a strong likelihood that interaction effects will be present.

Equation (11) gives the model for our 4 x 4 Latin Square design example and Table F-6

$$X_{ijk} = \mu + A_j + B_i + C_k + \epsilon_{ijk} \tag{11}$$

shows the arrangement for the ANOVA table. The critical region at an $\alpha = 0.05$ is shown in the right column as defined by the $F_{.05}$ with 3 and 6 degrees of freedom in numerator and denominator, respectively.

There are several variations of the basic Latin Square design. A Graeco-Latin Square design is one that has four factors each of which has each of its

TABLE F-6 - 4 x 4 LATIN SQUARE ANOVA EXAMPLE

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	$F_{.05}$
Between aircraft types, A_j	SS_a	4-1 = 3	$SS_a/3=MS_a$	MS_a/MS_ϵ	4.76
Between aircrews, B_i	SS_b	4-1 = 3	$SS_b/3=MS_b$	MS_b/MS_ϵ	4.76
Between missile groups, C_k	SS_c	4-1 = 3	$SS_c/3=MS_c$	MS_c/MS_ϵ	4.76
Experimental error, ϵ_{ijk}	SS_ϵ	16-9-1 = 6	$SS_\epsilon/6=MS_\epsilon$		
Totals	SS_T	15			

levels appear once and only once with each level of each other factor. Seldom is such a design useful since there are so few degrees of freedom for estimating the experimental error term, and because adding factors increases the number of interactions that must be assumed negligible.

A Latin Square design that doesn't have the same number of levels for each factor present in the design is called a Youden Square or incomplete Latin Square. Such a design may be appropriate if there is a logical reason for a fewer number of levels for a factor, (e.g., there are only three aircraft types of test while there are four groups of aircrews and missile production lot groups).

d. Factorial Designs. A factorial design is one that has all levels of a given factor combined with all levels of each other factor in the experiment. The factorial design is the most commonly used design in OT&E; the reasons for this popularity are brought out in the following discussion.

Consider an example of a factorial design which has the three factors that have been used in the previous examples: aircraft types, aircrews, and missile production lot groups. The test factor combinations are shown in Table F-7. There are two observations or replications in each cell in the matrix. By taking two or more replications per cell (or combination set of

test factors) the interaction effects between test factors may be tested and, also, separated out from the experimental error term. The example presented in the discussion of the two factor design is a factorial design (the Latin Square design may be regarded as a special case of a factorial design); however, there is only one observation per cell and based upon previous knowledge it was assumed that there was no interaction between aircraft types and aircrews. Hence, with this highly restrictive assumption, the previous two factor example problem is not a representative example of the general application of a factorial design.

TABLE F-7 - 3 FACTORS, 2 REPLICATIONS FACTORIAL DESIGN EXAMPLE

	Aircraft Types, A_j							
	I		II		III		IV	
Aircrews,	Missile Pd'n Lots, C_k		Missile Pd'n Lots		Missile Pd'n Lots		Missile Pd'n Lots	
B_i	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
1	X X	X X	X X	X X	X X	X X	X X	X X
2	X X	X X	X X	X X	X X	X X	X X	X X
3	X X	X X	X X	X X	X X	X X	X X	X X
4	X X	X X	X X	X X	X X	X X	X X	X X
5	X X	X X	X X	X X	X X	X X	X X	X X

The model for the factorial design example problem shown in Table F-7 is given by Equation (12)

$$X_{ijk\ell} = \mu + A_j + B_i + C_k + (AB)_{ij} + (AC)_{jk} + (BC)_{ik} + (ABC)_{ijk} + \epsilon_{\ell(ijk)}, \qquad (12)$$

where the terms in parenthesis represent the second and third order interactions between test factors and ℓ (ℓ= 1, 2) is the subscript that corresponds to the observation number within a cell. Table F-8 shows the ANOVA table for the example.

TABLE F-8 - FACTORIAL DESIGN EXAMPLE PROBLEM ANOVA TABLE

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	$F_{.05}$
Main Effects					
Between aircraft types, A_j	SS_a	4-1=3	$SS_a/3=MS_a$	MS_a/MS_ϵ	2.84
Between aircrews, B_i	SS_b	5-1=4	$SS_b/4=MS_b$	MS_b/MS_ϵ	2.61
Between missile groups, C_k	SS_c	2-1=1	$SS_c = MS_c$	MS_c/MS_ϵ	4.08
Interactions					
AxB	SS_{ab}	3x4=12	$SS_{ab}/12=MS_{ab}$	MS_{ab}/MS_ϵ	2.00
AxC	SS_{ac}	3x1=3	$SS_{ac}/3=MS_{ac}$	MS_{ac}/MS_ϵ	2.84
BxC	SS_{bc}	4x1=4	$SS_{bc}/4=MS_{bc}$	MS_{bc}/MS_ϵ	2.61
AxBxC	SS_{abc}	4x3x1=12	$SS_{abc}/12=MS_{abc}$	MS_{abc}/MS_ϵ	2.00
Experimental error, $\epsilon_{\ell(ijk)}$	SS_ϵ	80-39-1=40	$SS_\epsilon/20=MS_\epsilon$		
Totals	SS_T	79			

The principal reasons why the factorial design is used so often in OT&E are summarized:

(1) Generally, an OT&E test program has ambitious goals to investigate the effect of several test factors on the response variable. The factorial design permits the most efficient method to test several test factors (compared to a series of one-factor-at-a-time experiments).

(2) Every observation is used to estimate an effect from each test factor. Some level of each test factor is present in each observation.

(3) The experimental error is estimated over a wide range of test conditions and, generally, there is an adequate sample size (degrees of freedom) available for its estimation.

(4) When there are two or more observations per cell, an isolation and estimation of possible interaction effects between test factors may be performed.

e. Other Designs. There are numerous other ANOVA designs that may be applied to OT&E test programs. It is beyond the scope of this document to discuss the other designs in comprehensive detail. A brief mention is made below of some of the principal designs that are employed in OT&E. The test planner should consult with a statistician for the application of the proper statistical design.

A fractional factorial design is a factorial design that has an incomplete number of observations for at least one replication for the design matrix. For example, if one or more of the observations were not available for the factorial design shown in Table F-7, the design would be referred to as fractional factorial. Often, limited test resources and/or lack of interest for particular sets of combinations of test factors (cells) leads to fractional factorial test designs.

Whether a test factor is fixed or random will influence the form and interpretation of the F- tests that are made to test for significant effects. When all the levels of each factor in a design are fixed or set at pre-determined levels, the experiment has a fixed model. When all levels of each factor in a design are chosen at random, the test has a random model. When the design involves one or more factors that have their levels fixed and one or more factors that have their levels random, the experiment has a mixed model. In the example problem displayed in Table F-7, the aircraft type and missile production lot groups are fixed factors. If the aircrews are chosen at random, then the test has a mixed model.

There are test situations where there are test factors that are not factorial or crossed (taken in combination with) over all levels of each of the other test factors, (i.e., there is a test factor that is nested within a level of another test factor). For such a design, the experiment is called a nested experiment, (i.e., levels of one factor are nested within, or are subsamples of, levels of another factor). When an experiment involves (1) test factors that are crossed (factorial) with other test factors and (2) test factors that are nested within levels of other test factors, it is referred to a nested-factorial experiment

Table F-9 shows an example of the arrangement for a nested-factorial design. Each aircraft type has four missile launch racks. However, the same four launch racks are not used on each aircraft type. Thus, the launch rack effect, $C_{k(j)}$, is nested within the aircraft type factor. It is important to recognize when a design has a nested rather than a factorial test factor since the experiment model and ANOVA table breakdown is different compared to a completely factorial design.

There are many experimental situations where it is impractical to completely randomize the order of taking test observations among the levels of a test factor. For example, consider an experiment where there is interest in

TABLE F-9 - NESTED-FACTORIAL DESIGN EXAMPLE
RADIAL MISS-DISTANCE IS THE DEPENDENT VARIABLE

Aircrews, B_i	Aircraft Types, A_j															
	I				II				III				IV			
	Launch Rack, $C_{k(j)}$				Launch Rack				Launch Rack				Launch Rack			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2																
3																
4																

testing the performance (in terms of "time-to-target-acquisition") of three different radar types over four different target types and have three replications for each combination set of test factors. To achieve a completely randomized two factor design, a radar type would be used with one of the target types to be identified and then another radar type/target type combination would be selected at random for the next observation, and so on....... Such an experiment would entail flying for 36 separate combinations of radar type/target type conditions.

Fortunately, there is an ANOVA design called split-plot that can reduce the experimentation effort in terms of time and resources for this example problem. Table F-10 shows an arrangement for a split-plot design that could be used for this example. The nature of the test is such that an aircraft can fly to a target type and there could be a measure of the performance for each of the three radar types. Next, a second target type could be displayed in place of the previous target type and the aircraft flown over the target so that the three radar types could be measured for their performance in detecting the target. This process would be continued over the four target

TABLE F-10 - SPLIT-PLOT DESIGN EXAMPLE

Time-To-Target Acquisition Is the Dependent Variable

Replication, B_i	Radar Type, C_k	Target Types ("Plots"), A_j			
		I	II	III	IV
1	C_1				
	C_2				
	C_3				
2	C_1				
	C_2				
	C_3				
3	C_1				
	C_2				
	C_3				

"Split Plots"

"Whole Plots"

types (in a random order) and then repeated for the second and third replication. Thus, a total of 12 sorties would be required rather than 36 as in the completely randomized factorial design.

The four target types are referred to as plots. ("Plots" comes from terminology used in agricultural ANOVA applications). A main effect in the test, target types, is confounded with plots. It is impossible to tell the difference between an effect caused by a plot from an effect caused by a target type.

The "replication by target type" cells (this includes 12 cells as indicated in Table F-10) are called whole-plots. The "radar type by target type" cells (there are 3 cells as indicated in Table 10) are called split plots. In the example one main effect (target types) is confounded with plots and the other main effect (radar types) is not. Hence, for this problem and as a general

rule, it is desirable to place the test factor of most concern as the factor which is not confounded with plots. The ANOVA F-tests that are performed to identify significant effects in the split-plot design utilize appropriate interaction terms as an estimate of the experimental error.

General Linear Hypothesis Model. A convenient method for computing and solving ANOVA problems for OT&E test programs is the application of the general linear hypothesis model. It frequently happens in operational tests that there are unbalanced designs (due to resource limitations and/or lack of interest in particular combinations of test factors) and missing data for observations that were originally planned for in the test design. The application of the general linear hypothesis model to analyze the test data permits a maximum utilization of the available data in the presence of unbalanced designs and missing observations for the design cells.

After ANOVA. After the ANOVA F-tests have been performed, attention may be centered on the determination of the manner in which the significant effects occur. To illustrate, if the aircraft type is a significant effect in one of the previously discussed example problems, then the questions arise as to "Which aircraft types are significantly different from each other?" "Which groups of aircraft types are significantly different from other groups of aircraft types?"

If a decision is made prior to the conduct of the experiment as to which combinations of means to compare to each other, the method of orthogonal contrasts may be used with no change in the Type I error risk α. The method of orthogonal contrasts is performed after the ANOVA and has several restrictions as to which combinations of comparisons can be made.

If the decision for which comparisons are to be made is delayed until after the experimental data may be examined, comparisons can still be made, but the Type I error α risk is changed since such decisions are not taken at random but are based on observed results. There are several methods for performing the tests on means. One of the easiest techniques to apply is Duncan's Multiple Range Test.

Another item of interest that may be estimated after the ANOVA tests is the percent of variability in the dependent variable that is due to a significant effect and that which is attributable to the experimental error. The test experimental design model and parameters that have already been computed for the ANOVA F-tests are utilized to estimate the percent contributions to variability due to the significant test factors.

In the discussion of the various ANOVA models, mention was made to and a definition given for interaction effects between two or more test factors. To visualize graphically how an interaction would look consider Figure F-1 which shows a plot for the aircrew and aircraft type combination for the previously discussed 3 factor, 2 replication factorial design. There is no interaction between aircraft types I, II, and IV with aircrews. Notice that their connecting lines to the plotted points are approximately parallel. There is a significant interaction between aircraft type III, the other three aircraft types, and aircrews. The interaction shows up graphically by the non-parallel (and crossing) connecting lines between the plotted points. The change in response

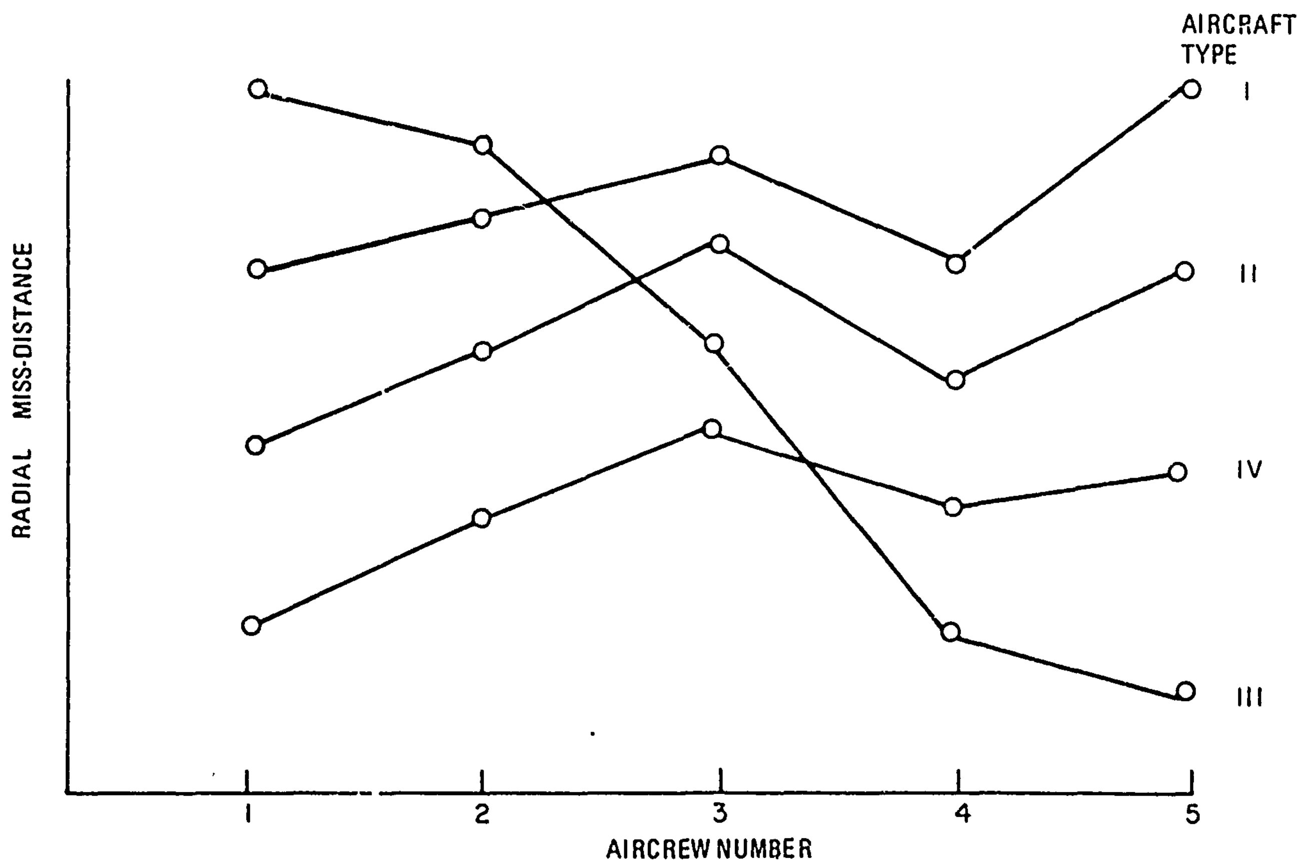


Figure F-1. Two factor example problem interaction effect

(radial miss-distance) between levels of one factor (aircrews) is not the same for all levels of the other factor (aircraft types).

The previously discussed ANOVA example problems have had test factors that are qualitative, (e.g., aircraft types, aircrews, missile production lot groups). In a factorial design if these are factors that **are quantitative, (e.g., weapon release altitude)** and equally-spaced, (e.g., 10,000, 20,000, 30,000, and 40,000 ft) then more information can be analyzed from the test data. Using orthogonal polynomial coefficients, test factors may be tested for significant effects that are related to the manner in which the dependent variable is affected by the factor. For example, the relationship between the independent and dependent variables could be linear, quadratic, or cubic. The orthogonal polynomials are used in conjunction with the ANOVA F-test to determine significant polynomial effects for equally-spaced quantitative test factors.

a. Comparing More Than Two Variances. One of the basic assumptions in ANOVA models is that the experimental error for each cell in the statistical design matrix is homogeneous, (i.e., σ_ϵ^2 is the same in each cell). Moderate departures from this assumption do not seriously affect the validity of the F statistics. Stated another way, the ANOVA F-test is robust with regard to the assumption that the design cell experimental error is homogeneous.

   There are several tests available to test for homogeneity of variance. The most commonly used test is Bartlett's test. The routine use of Bartlett's test is not recommended due to relatively complex calculations and several highly restrictive assumptions that are required. There is a homogeneity of variance test proposed by Hartley and another one proposed by Cochran that are relatively simple to apply and give comparable results to Bartlett's test.

b. Determining Relationships. The situation can arise in OT&E where there is interest in determining the relationship (if it exists) between a dependent variable and one or more independent variables. The available data for the analysis may be from a test program and historical data files; however,

quite often the test data was not originally planned for use in determining the relationships of current interest. The principal analysis technique used to examine relationships is called multiple regression analysis.

As an illustration, consider the simple case where it is desired to estimate miss-distance (dependent variable) as a function of range (independent variable). Such test data, when plotted, may look similar to Figure F-2.

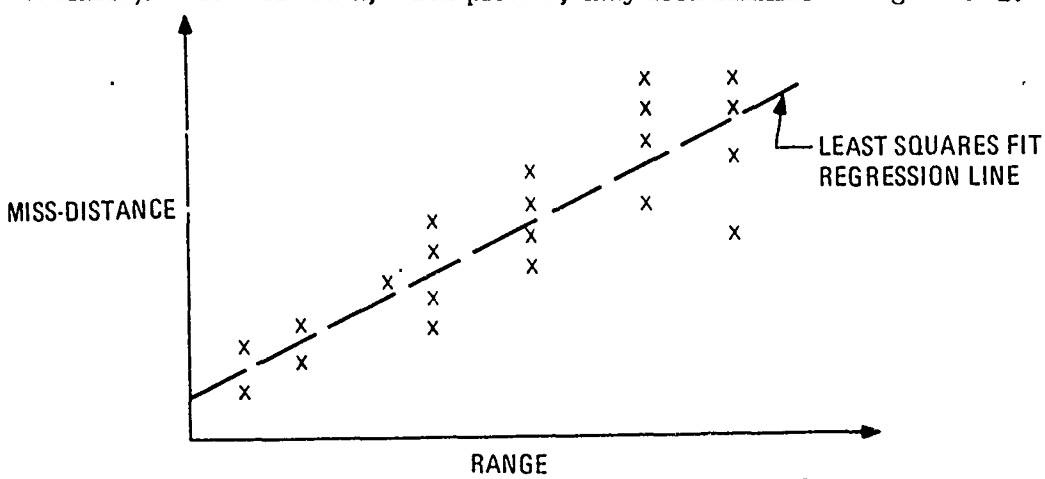


Figure F-2. Plotting the relationship of variables

By subjecting the coordinate points from the test data to a multiple regression analysis, the equation for a linear function (dash-line in sketch) may be derived. Since it is linear, it will be of the form $y = a + bx$ where the analysis technique solves for the coefficients a and b. The solution minimizes the sum of the square of the vertical distance between the data points and the regression line (least squares fit).

It is entirely possible (and frequently occurs) that there is no significant relationship between the dependent and independent variables. Stated another way, knowing a particular value for the independent variable will be of no assistance in predicting the value of the dependent variable. In the multiple regression analysis, a t-test is performed to determine if the independent variable has a significant effect on the dependent variable.

The fact that the model for the multiple regression analysis is linear does not restrict its use to analyzing only linear functions. By the use of transformation of variables, a new variable that is linear can be used to replace one that is nonlinear, (e.g., e^x replaced by z). For example, a particular analysis may require determination of the relationship between weapon accuracy (y) and aircraft velocity (x) using experimental data. There are a large number of possible functional relationships to examine in order to find the best one. To illustrate a few, consider the five candidates below:

$$y = a + bx$$

$$y = a + bx + cx^2$$

$$y = a + cx^2$$

$$y = a + bx + dx^3$$

$$y = a + bx + cx^2 + dx^3$$

Computationally, it is desirable to have an efficient method to search through the candidate relationships to select the best one. Stepwise Multiple Linear Regression will accomplish this efficient search by the following scheme:

(1) The input to the problem contains the definition of the candidate variables (x, x^2, and x^3 in the example).

(2) One variable is added to the regression equation at a time. The variable added is the one that produces the greatest reduction in the residual error term.

(3) As an option, variables can be forced into the regression.

(4) In the iterative procedure, non-forced variables are removed from the equation if their new (recomputed) contribution to the reduction of the error terms becomes too low.

(5) Optionally, the regression equation may or may not be required to have an intercept ("a" in the example).

Non-Parametric Techniques

When the characteristics of the probability distribution for the response (dependent) variable are unknown or ignored, the use of nonparametric (or distribution free) statistics may be appropriate. "Characteristics" of the probability distribution refers to such things as the type of distribution, (e.g., Normally or uniformly distributed) and the distribution parameters, (e.g., mean and variance). Non-parametric techniques are general in that they are independent of any assumption of the characteristics of the probability distribution of current interest; however, the techniques are frequently used to test the hypothesis that a specific sample of observations comes from a population with a particular type of distribution, (e.g., Normally distributed).

Compared to parametric techniques, non-parametric techniques do not glean as much mileage out of the same amount of data, (i.e., non-parametric techniques are not as powerful or discriminating). Despite the reduction in "power", non-parametric techniques can be very useful and sometimes they offer the only remaining course of action. Typical applications of non-parametric techniques include: (1) goodness of fit tests, (2) 2x2 contingency table, and (3) related sample tests.

Goodness of Fit Tests

There are many parametric techniques that require the assumption of a particular type of probability distribution for a variable. One of the most common assumptions is that the variable be distributed Normally. A goodness of fit test is a statistical test to provide reassurance that the distribution assumption is reasonably correct based upon the sample data from the population; and responds to the question, "Does the data fit the assumed probability distribution function?"

Generally, goodness of fit tests are based upon the differences between "sample probability distribution function" and an assumed probability distribution function.

Individual goodness of fit tests vary in the manner in which they measure the differences. Two common goodness of fit tests are the chi-square and the Kolmogorov-Smirnov test.

a. Chi-Square Test. The oldest and most frequently used in goodness of fit test is the chi-square test. The test is only valid for large sample sizes; however, it is applicable to either discrete or continuous distributions and for testing partially or completely specified distributions.

The essence of the chi-square test is that a sample of n observations of x_i ($i = 1, 2, \ldots, n$) is sub-divided into J adjacent class intervals. The frequency of observations, f_j, in each of the intervals is compared to the number of such observations that one would expect to find in the interval for the assumed distribution. If P_j is the probability that an observation from the assumed distribution would lie in the j^{th} interval, then the expected value of f_j is np_j. Thus, the chi-square statistic is computed by Equation (14)

$$\chi^2 = \sum_{j=1}^{J} \frac{[f_j - np_j]^2}{np_j} \qquad (14)$$

For large samples, χ^2 has approximately the chi-square distribution with J-1 degrees of freedom. The critical region for rejecting the null hypothesis (H_0: the sample distribution comes from a population whose distribution is the same as that assumed in the test) is $\chi^2 > K$, where K is found in tables for the chi-square statistic and J-1 degrees of freedom.

As a guideline to ensure that the sample size is large enough for the proper application of the chi-square test, each of the J intervals should have $np_j \geq 5$. (This applies to both the discrete and continuous distribution cases.) If the sample data is going to be used to estimate parameters, (e.g., the mean and variance) for the assumed distribution that is used in the chi-square test, one degree of freedom must be subtracted for each such parameter from the total number of degrees of freedom available for the hypothesis test. For example,

in testing for Normality where the sample mean and variance are used in the assumed distribution, the number of degrees of freedom is J-1-2 = J-3.

b. Kolmogorov - Smirnov Test. The Kolmogorov-Smirnov Test (hereafter denoted K-S for brevity) has the advantage that it is applicable for small as well as large samples. The test statistic for the K-S test is given by Equation (15)

$$D_n = \max_{-\infty < x < \infty} \left| F_n(x) - F(x) \right| , \tag{15}$$

where $F_n(x)$ is the sample cumulative probability (distribution) function and $F(x)$ is the cumulative probability distribution function for the population that is under test. Figure F-3 shows a hypothetical example of the graphical interpretation of the meaning of Equation (15). D_n is the maximum difference (vertical distance in Figure F-3) which exists between the two distribution functions.

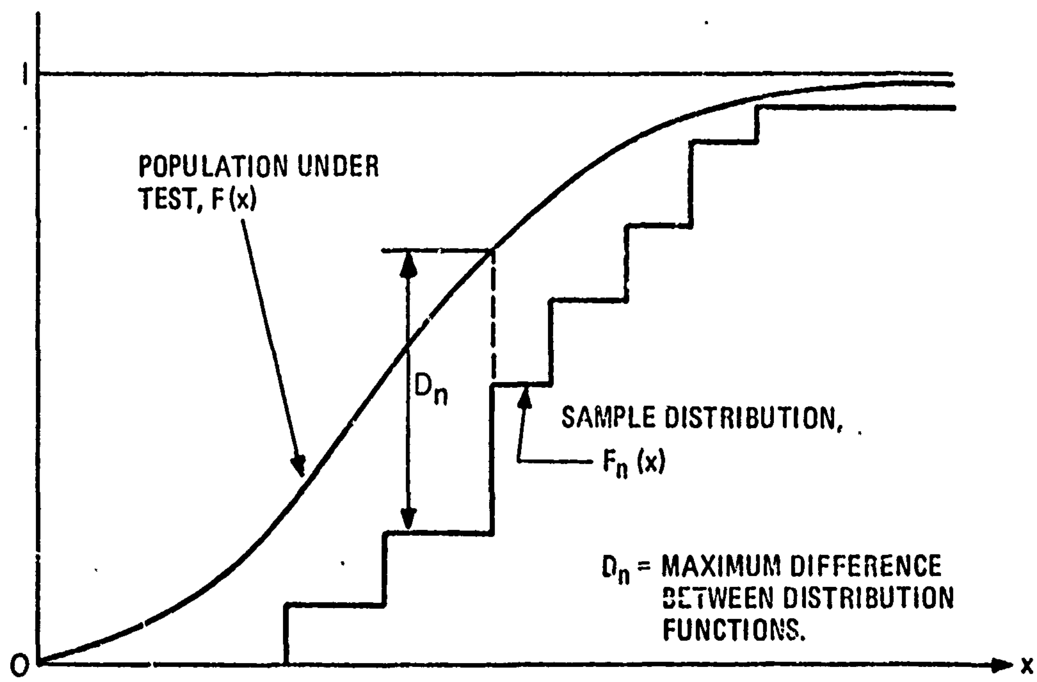


Figure F-3. Kolmogorov-Smirnov Statistics

The sample statistic D_n is independent of F (X) and there are table values (K) for large and small sample sizes that define the critical region for rejecting the null hypothesis (H_o: the sample distribution function is the same as the population distribution function which is under test), i.e., $D_n > K$.

The K-S test is based upon a continuous distribution function for the sample statistic; however, the test may be used with discrete distributions since such applications will be on the conservative side. The K-S test is applicable to experiments where there is a natural way to assign numerical values to the outcomes of the random variable of interest, x. For situations where numerical values are arbitrarily assigned to the outcome of chance events, (e.g., 0 to a head and 1 to a tail in a coin flip), the K-S test is not applicable since the results of the test may be changed by simply reassigning a different set of numerical values to the event outcomes, (e.g., 1 to a head and 0 to a tail in a coin flip).

2 x 2 Contingency Table. The Fisher exact probability test enables user to test, at an exact confidence level, the hypothesis that two independent, large and unequal samples of dichotomous data are drawn from the same population. The test is used when the scores from two independent groups of random samples fall into one or the other of two mutually exclusive classes. Hence, each experimental unit in both sample groups obtains one of two possible scores. The scores are represented by frequencies in a 2 x 2 contingency table. To illustrate, consider the contingency table below where subjects from two groups are scored with either a "success" or "failure" for a mission.

	Number of "Successes"	Number of "Failures"	Row Total
Group 1	17	36	53
Group 2	23	47	70
Column Total	40	83	

The Fisher test is used to determine the probability that the two groups differ significantly in the proportion with which they fall into the success and failure classification. The model for the test considers the marginal totals as fixed, (e.g., 17+36, 23+47, 17+23, 36+47) and then utilizes the properties of the hypergeometric distribution to compute the probability of occurrence (P_0) of the four numbers in the table or cases more extreme.

Since the data is discrete (integers), a special provision is necessary to test the hypothesis of interest, (i.e., H_0: Groups 1 and 2 come from the same population) at exactly a given confidence level. This is accomplished by employing Tocher's modification which entails selecting a random number from a uniform distribution. Having defined a Type I error risk of α, the critical region for the hypothesis test is for values of $P_0 < \alpha$.

Two Related Samples. The situation can often arise where the experimenter wishes to compare two phenomena (experimental conditions) by "controlling" the test subjects to perform or exercise each of the two test conditions. One way to "control" the test is to have each subject serve as his own control by performing under both conditions at different times. If it is illogical to have a test subject perform both test conditions of interest, another alternative way to control the test is to have matched pairs of the test subjects (nearly identical as possible) and have each one of the pairs perform only one of the test conditions of interest, another alternative way to control the test is to have matched pairs of the test subjects (nearly identical as possible) and have each one of the pairs perform only one of the test conditions of interest. The following subsections discuss two common tests for testing two related samples.

a. Sign Test. The sign test may be employed to test where there is a significant difference in the two experimental conditions from which related samples of experimental subjects are being drawn. The test derives its name from the fact that it uses plus (+) and minus (-) signs rather than quantitative measurements as its data. It is particularly useful in experimentation when quantitative measurement is impractical or infeasible; however, it is possible to rank members of

related samples with respect to each other for each experimental condition, e.g., aircrews that fire two missiles in training are more effective than aircrews that fire only one. The test is nonparametric: no assumptions need to be made with regard to the form of the distribution of differences, nor does one need to assume that all subjects are drawn from the same population. For example, the different pairs may be from different populations with respect to age, sex, intelligence, etc., the only requirement is that within each pair the experimenter has achieved matching with respect to the relevant variables.

To illustrate a situation where the sign test is useful, consider the table below which has "hit" or "miss" indications for aircrews corresponding to their first and second shot at a given target. A plus (+) and a minus (-) sign is recorded to denote an improvement or degradation in each pair, respectively. The sign test is used to test the hypothesis that there is no difference between the performance for aircrews between the first and second shot. If there is no difference, one would expect there to be the same number of each type of sign. The hypothesis is rejected at a given confidence level if too few differences of one sign occur.

AIRCREW TEAM NUMBER

	1	2	3	4	5		7	8	9	10	11	12
First Shot	M	H	H	H	M	M	M	M	H	M	M	M
Second Shot	H	M	H	M	H	M	H	H	M	H	H	H
Improvement Sign	+	-		-	+		+	+	-	+	+	+

M - Miss
H - Hit

The basis for the sign test is that if the first shot performance is the same as the second shot performance then the probability that the performance for shot 1 for an aircrew is greater than the performance for shot 2 is equal to 1/2. Also, for the reverse situation, the probability is 1/2. Symbolically, this may be represented by Equation (16)

$$P(x_1 < x_2) = P(x_1 > x_2) = 1/2, \tag{16}$$

where x_i (i=1, 2) is the indication of the performance for trial i. If the performance of each of the two shots are equal, then we would expect the number of times $x_1 > x_2$ to equal the number of times $x_1 < x_2$. The test for the hypothesis, H_0: that the two shots are equal, rejects H_0 if too few of one of the inequalities occurs compared to the occurrences of the other.

Let k_1 = the number of times $x_1 > x_2$ occurs,

and k_2 = the number of times $x_1 < x_2$ occurs,

$n = k_1 + k_2$,

$k = \min(k_1, k_2)$,

then the probability, P, of getting k (or more extreme) occurrences out of n is given by the binomial distribution

$$P = \sum_{i=0}^{k} \binom{n}{i} p^{i} (1-p)^{n-i}, \text{ where } p = 1/2 \tag{17}$$

After chosing a Type 1 error risk level α, the hypothesis H_0 is rejected if $P \leq \alpha$. For the aircrew example, the expression corresponding to Equation (17) is given below.

$$P = \sum_{i=0}^{3} \binom{10}{i} \left(\frac{1}{2}\right)^{i} \left(\frac{1}{2}\right)^{10-i} = 0.172$$

For an $\alpha = 0.10$, the hypothesis H_0 would not be rejected.

As was the case with Fisher's exact probability test, the sign test has a discontinuous variate (integers). Therefore, Tocher's modification has application to permit one to test at an exact confidence level. The modification entails selecting a random number from a uniform distribution.

b. Wilcoxon Matched Pair Test. The previously discussed sign test utilizes data concerning the direction of change between the two test conditions of interest. A more powerful test, the Wilcoxon matched pair test, may be employed if not only the direction, but the relative magnitude of the direction difference is estimated. A matched pair which has a large difference between the responses for the two test conditions is given more weight than a pair which shows a small difference.

Quite often the precise measurement of the performance of a test subject for a test condition is not feasible; however, a score which is assigned to it and other such trials is useful for obtaining the relative differences and ranking them for comparison purposes. If in the example problem for the sign test an estimate for the miss-distance was made (instead of simply the designation "hit" or "miss"), the Wilcoxon matched pair test could have been employed with an increase in discriminating power to detect differences in experimental (test) conditions.

5. DEVELOP AND ANALYZE ALTERNATIVE STATISTICAL DESIGNS

Having identified the general type of statistical technique that would have application towards solving the particular test problem, the experimentor should not identify alternative statistical designs (variations of the basic statistical design technique selected) and corresponding sample/replication sizes to help determine the best design for the test. Since some of the factors affecting an OT&E program are military judgmental in nature (rather than being strictly statistical), the alternative statistical designs for the test should be displayed along with a discussion of their advantages and disadvantages so that decision-makers can perform a convenient review.

Identifying Alternative Statistical Designs

There are a number of reasons why the experimenter should consider alternative statistical designs for the test before recommending a specific design. Included in the list of reasons which could dictate variations in the design are the following (the discussion is oriented toward the use of parametric techniques):

a. Whether or not there may be significant interaction effects.

b. The number of levels of each factor to include in the test.

c. The combinations of levels of two or more factors to include in the test.

d. The training requirements for test subjects.

e. The sample/replication size requirements.

f. The alternative results which may arise from exploratory trials (before the main experiment).

As an illustration of the alternative statistical designs which could arise for a test, consider the following example. There are two weapons to compare. In one basic design, aircrews could fire one and only one of the weapons to minimize the training requirements. In another basic design, the aircrews could fire both weapons to hedge against the situation that one of the weapon-aircrew groups had a higher caliber of personnel. A third basic design factor could be whether or not the aircrews perform more than one replication per weapon. Replicating permits the analysis of interaction effects; however, it may also introduce "learning" effects from having repeated a previous trial event.

Table F-11 shows a display for six alternative statistical designs which are derived for the example problem. The following notation is used in the table:

y_{ijk} = the value for the dependent variable,

n = number of aircrews in a design,

μ = the population mean,

W_i = the weapon effect,

A_j = the aircrew effect,

R_k = the replication effect, and

ϵ_{ijk} = the random error.

TABLE F-11. SIX ALTERNATE STATISTICAL DESIGNS FOR THE EXAMPLE PROBLEM

Case 1 - Aircrews Use Only One Weapon

Design 1. One replication per aircrew.

Design Matrix

Weapons

Aircrews		W_1		W_2
	1		2	
	3		4	
	o		o	
	o		o	
	o		o	
	n-1		n	

Model

$$y_{ij} = \mu + W_i + \epsilon_{ij}$$

Assumptions

1. No interactions.
2. No aircrew differences.

Design 2. Two or more replications per aircrew.

Design Matrix

Weapons

Aircrews		W_1		W_2
	1	✓	2	✓
	3	✓	4	✓
	o		o	
	o		o	
	o		o	
	n-1	✓	n	✓

Model

$$y_{ijk} = \mu + W_i + A_j\ (i) + R_k + E_{ijk}$$

Assumptions

1. No interactions.
2. Possibility of significant replication effect.

TABLE F-11. SIX ALTERNATE STATISTICAL DESIGNS FOR THE EXAMPLE PROBLEM (Continued)

Design 3. Two or more replications per aircrew.

Design Matrix

Same as for Design 2.

Model

$$y_{ijk} = \mu + W_i + A_{j(i)} + \epsilon_{ijk}$$

Assumptions

1. No interactions.
2. No significant replication effect.

Case 2. Aircrews Use Both Weapons

Design 4. One replication per test condition.

Design Matrix

Weapons

Aircrews	W_1	W_2
1		
2		
o		
o		
o		
n		

Model

$$y_{ij} = \mu + W_i + A_j + \epsilon_{ij}$$

Assumption

1. No interactions.

Design 5. Two or more replications per test condition.

Design Matrix

Weapons

	W_1	W_2
1	√√	√√
2	√√	√√
o		
o		
o		
n	√√	√√

Model

$$y_{ijk} = \mu + W_i + A_j + R_k + (WA)_{ij} + (WR)_{ik} + (AR)_{jk} + \epsilon_{ijk}$$

Assumptions

1. Possibility of significant replication effect.
2. Possibility of significant interaction effects.

TABLE F-11. SIX ALTERNATE STATISTICAL DESIGNS FOR THE EXAMPLE PROBLEM (Continued)

Design 6. Two or more replications per test condition.

Design Matrix	Model
Same as for Design 5.	$y_{ijk} = \mu + W_i + A_j + (WA)_{ij} + \epsilon_{ijk}$

Assumptions

1. No replication effect.
2. No replication interaction effect.

The designs for the example problem displayed in Table F-11 demonstrate how, for a relatively few number of test factors, six alternative models were developed that should receive individual consideration before a specific design is recommended. If exploratory trials were conducted prior to the actual test, insight might be obtained as to whether there is a significant replication effect and this could influence the selection of the appropriate model, (e.g., if preliminary trials indicate a significant replication effect, models 2 and 5 become leading contenders).

Determining the Number of Observations per Design

One of the most important steps in analyzing alternative designs for the test is the determination of the proper number of observations (sample size and number of replications or repeated trials with the same test subjects) to be taken during the test. The sample, replication size affects the test program requirements for resources, trials, schedule, and personnel training.

When one is interested in estimating the mean of a Normally distributed random variable, a measure of the probable accuracy of the estimation procedure is the expected confidence interval length, (ECIL). The ECIL is a function of (1) the confidence with which one would like to be sure that the interval contains the true population mean or $1-\alpha$, (2) the number of samples n, (3) the number of replications per sample r, (4) the assumed proportion (k) of the total variability that is due to between test subjects as compared to within a test subject's performance, and (5) the population standard deviation σ. For a particular test, one may vary α, n, r, and k and examine the ECIL in terms of units of σ. Thus after fixing α and k the relative magnitude of the ECIL may be determined as a function of combination sample-replication sizes.

Figure F-4 shows a plot of the ECIL for four confidence levels $(1-\alpha)$ and for the case where there is one replication per sample **(r=1)**. For this example, notice that the curves begin to flatten out for $n \geq 10$ and, thus, additional sampling has a lower marginal value in terms of the relative scale of measurement σ.

When one is concerned with detecting a specified difference between two means, the sample size requirements are a function of (1) the decision risks (α and β) that one is willing to take and (2) the magnitude of the difference between the two means relative to the dispersion or variability in the response variable.

Let us review the implication of the α and β error risks in terms of comparing two weapon systems, i.e., a proposed new system versus an existing weapon system). α is the probability of concluding that the new system **is** better than the old system when if in fact there is no difference. A mistake of this kind may lead to a cost in

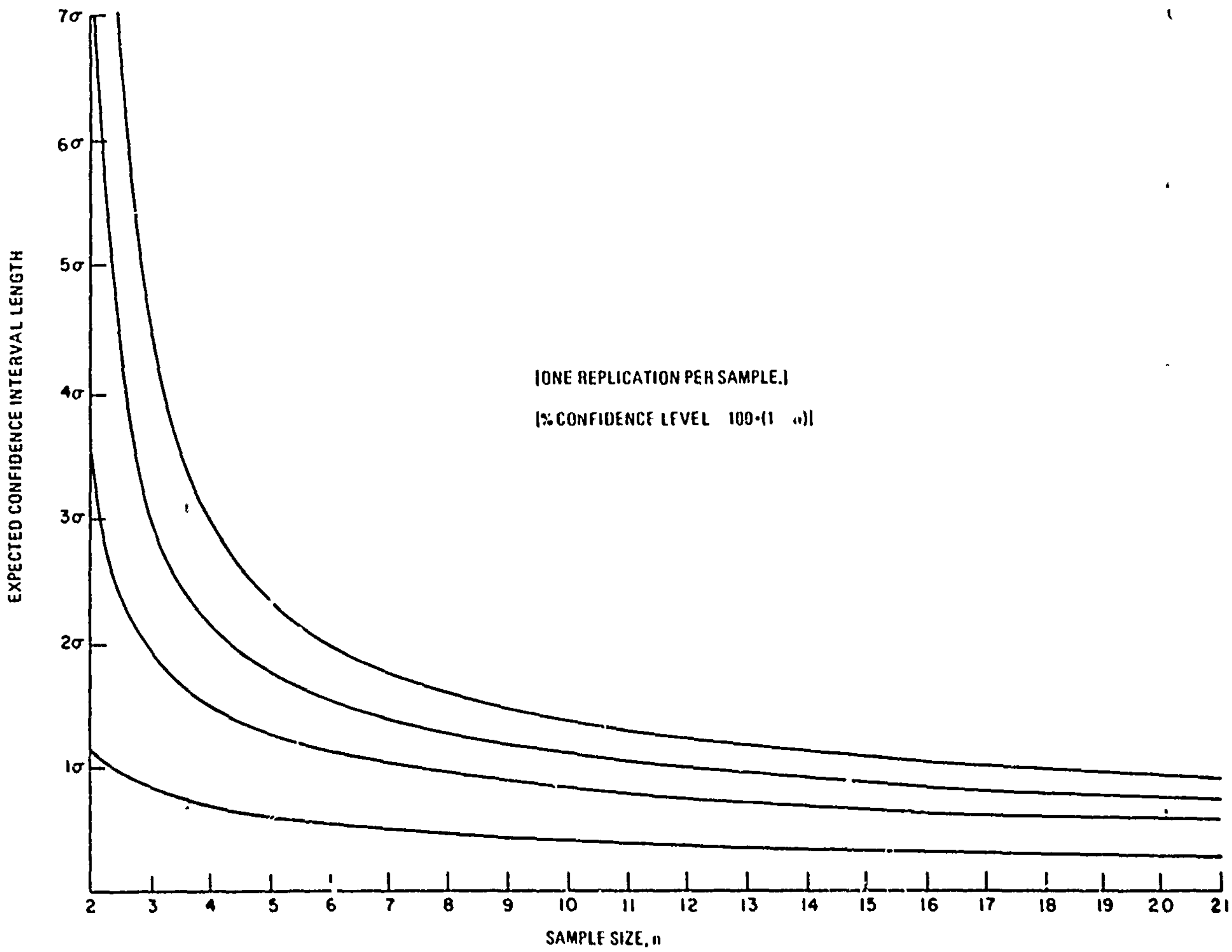


Figure F-4. Expected confidence interval

dollars if the new system is acquired; however, there will be no loss in system performance. As one would anticipate, if α and β are made small, a large sample size is required; if α and β are allowed to be large, a smaller (in relative terms) sample size is required.

The typical manner in which the true difference (Δ) in system mean performance that is desirable to detect is expressed in units of the standard deviation σ. If the ratio Δ/σ is small, a large sample is required and vice versa. In general, the value of σ and Δ have to be estimated from historical data taken under similar test conditions or from the results of exploratory trials.

Figure F-5 shows a plot of the ratio Δ/σ versus sample size for $\beta = 0.10$ and $\alpha = 0.025$, 0.05, 0.10, 0.15, and 0.20. One can see that the number of observations required is very sensitive to the Δ/σ ratio. Hence, when resources are expensive or in short supply the α, β, and Δ are clearly important parameters in the test program,

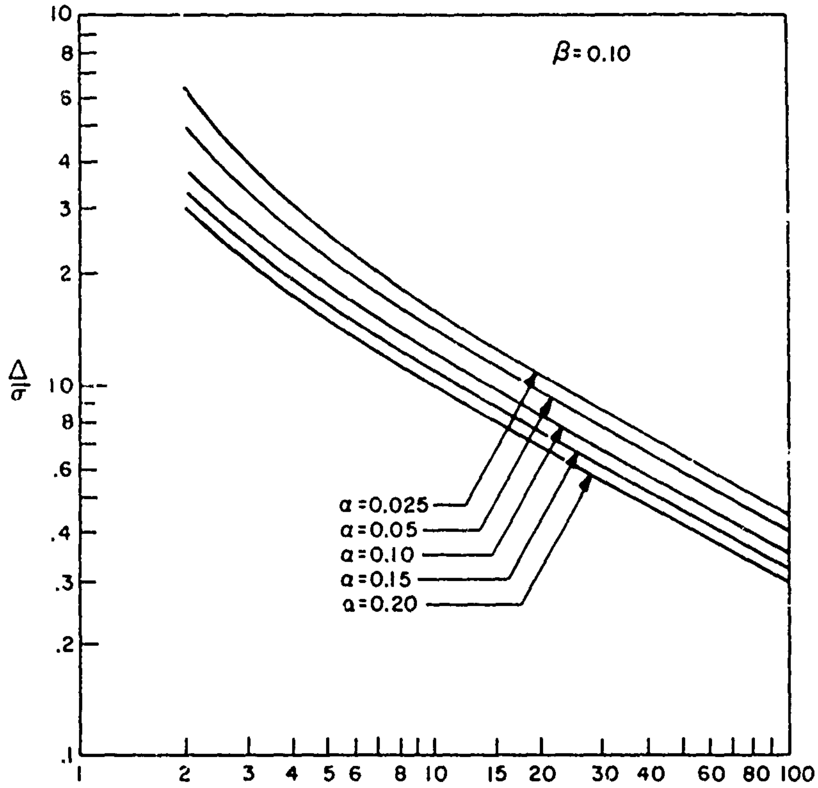


Figure F-5. Number of observations needed for comparison of two means - one-sided test

a reasonably good estimate of Δ/σ should be made to ensure that test results will be responsive to the test objectives.

It should be noted that a relatively few number of samples are required from exploratory trials, (e.g., 10) to estimate Δ/σ for an experiment that may require a large sample, (e.g., 80). The reason for this lies in the fact that if one goes back to Figure F-4, you will remember that the expected confidence interval length is in units of σ. For a given population, σ is fixed and there is nothing that can be done to change it. Despite a large confidence interval when σ is large, the percentage decrease in the length of the confidence interval caused by an increase in the sample size does not depend on the actual value of σ. Thus, the practical strategy is to simply take enough samples to place one on the flat part of the curve, (e.g., $n \geq 10$). Exploratory trials also offer the opportunity to check test procedures and identify

illogical scenarios that would make system comparisons difficult, (e.g., target saturations that offer little room for improvement for new systems).

For test designs that utilize analysis of variance models, the power of the F-test must be examined to determine the sample size requirements. For test designs for which non-parametric techniques are to be employed, there are tables and charts available in the statistical literature to help guide the test planner in choosing an adequate sample size.

Performing Design Tradeoffs

Having identified the alternative statistical designs for the test, the experimentor should next examine the relative advantages and disadvantages of each design and perform tradeoffs analyses with any parameters which are subject to optimization.

As an example of a design tradeoff analysis, consider a test where there is interest in estimating the average value μ for an aircrew's performance (which is some Normally distributed random variable) for using a new weapon system. There will be aircrews replicated r times to estimate μ. There is c dollars available for the experiment. It costs c_1 dollars to train an aircrew for the test and c_2 dollars to make an observation per replication on any given sample. Therefore, the problem is to determine the values for n and r which will provide the best estimate for μ subject to the constraint

$$nc_1 + nrc_2 \leq c. \tag{18}$$

The best estimate for μ is an adaptation of Equation (1) and is given below by Equation (19)

$$\overline{X} = \frac{1}{nr} \sum_{i=1}^{n} \sum_{j=1}^{r} X_{ij}. \tag{19}$$

The estimate for the variance of μ is given by Equation (20)

$$S_x^2 = \frac{1}{n} (\sigma_b^2 + \frac{1}{r} \sigma_w^2). \tag{20}$$

where σ_b^2 is the variance between aircrews and σ_w^2 is the variance within an aircrew. Using standard techniques of calculus it can be shown that the optimum values of n and r (n^* and r^*) are given by Equations (21) and (22), respectively.

$$n^* = (c/c_2) \Big/ \left[c_1/c_2 + (\sigma_w/\sigma_b) \sqrt{c_1/c_2} \right] \tag{21}$$

$$r^* = \max \left[(\sigma_w/\sigma_b) \quad c_1/c_2,\ 1 \right] \tag{22}$$

Since the values for n^* and r^* as computed by Equations (21) and (22) may not be integers it may be necessary to perform (at most) two calculations to get the optimal integer solution. Let $[r^*]$ and $[n^*]$ be the largest integers less than or equal to r^* and n^*, respectively. Then the variance of $\overline{X}$ using $[n^*]$ samples and $\left[\frac{c - [n^*]\, c_1}{c_2\, [n^*]}\right]$ replications must be computed and compared with the variance of $\overline{X}$ using $\left[\frac{c}{c_1 + c_2\, [r^*]}\right]$ samples and $[r^*]$ replications; whichever of these two methods gives (n, r) combination that minimizes the variance of $\overline{X}$ is the solution. (It may be that both of the above calculations yield the same (n, r) combination and in this case the same pair is the obvious solution.)

Another area for consideration for design tradeoffs is with regard to the requirements for instrumentation measurement error (when recording the dependent variable). A Test Director may ask the question, "Should I have measurements made to the nearest 35 meters? 5 meters? yard? foot? or inch?" The increased variable from measurement error will add to the sample size requirements. A tradeoff can be made between the cost for reducing the measurement error and the cost for additional samples and replications.

Recommending a Statistical Design

The alternative statistical designs should be displayed for convenient review by decision-makers along with a discussion of their advantages and disadvantages. The display for each design should include:

a. Design Matrix
b. Math Model
c. ANOVA Table
d. Assumptions
e. Sample and Replication Sizes
f. Resources for Training and the Experimental Trials

The statistician supporting the test program should identify his recommendation for a test-design(s). His recommendation may involve the use of either of two designs depending upon the outcome of the results of exploratory trials, (e.g., whether or not there is a significant replication effect). The actual design selected for the test will come from the office of the Test Director.

6. IDENTIFY CONDUCT OF THE SELECTED TEST DESIGN

The assumptions accompanying the structure of the math model for the statistical design that is selected for the test will dictate the randomization and control requirements for the test trials.

Method of Randomization

The proper randomization within a test design is intended to spread the random effects of extraneous factors, (i.e., confounded into the random error term) over as many of the levels of test factors as possible to reduce the possibility that an erroneous bias could be introduced in the data. The method of randomization could be the assignment of trial numbers to test subjects based upon the selection of random numbers or the random selection of a test design matrix trial table and the subsequent assignment of test subjects to the design matrix cells.

Testing Order, Schedule, and Control

Controlling an OT&E experiment for the proper test conditions is noticeably more difficult than controlling an experiment in a laboratory. The statistician supporting the test program should take into account the practical operational limitations for scheduling test subjects for experiment trials and devise an appropriate test trial order with a corresponding test schedule for test subjects. Relative to each test trial there is the requirement for test control which corresponds to the levels of the test factors that are present in each design matrix cell. Extraneous factors which are not in the test design should be "controlled" to a practical extent possible so that the test data is not unduly contaminated by their presence.

7. DEVELOP DATA ANALYSIS PLAN

Basically, the statistical design for the test dictates (defines) how the test data will be analyzed. However, there is benefit in specifically identifying the exact steps in the data analysis plan so that (1) previously overlooked contingencies in data results may be planned for, (2) computer programs to automate the calculations may be obtained, and (3) the results from the statistical analysis may be planned for convenient interfacing with the evaluation process that is keyed to respond to the basic test objectives and questions.

Computing Test Statistics

This section includes a discussion of (1) a procedure that can assist in treatment of outliers in the data, (2) transformation of variables that may assist in satisfying Normality requirements in the dependent variable, and (3) statistical analysis computer programs that can help automatic computations.

Treatment of Outliers

It is often vague or ambiguous in an experiment as to how extreme values (or outliers) will be identified and treated in test data. The reasons for the omission or inclusion of an outlier may be based on operational and/or statistical considerations. The purpose of the discussion here is to present the general philosophy and a sequence of steps to assist in an orderly treatment of outliers.

An outlying observation or "outlier" is one that appears to deviate markedly from other members of the sample in which it occurs. An outlier may be merely an extreme manifestation of the random variability inherent in the population from which the data was taken. If this is true, the value should be retained and processed in the same manner as the other observations in the sample. Other possibilities are that the extreme value may be the result of a gross deviation from prescribed experimental procedure or an error in recording or calculating the numerical value. In such cases, it may be desirable to institute an investigation to ascertain the reason for the aberrant value. The observation may eventually be rejected as a result of the investigation, though not necessarily so.

If an outlier is known to have occurred due to a gross deviation from prescribed experimental procedure, it should be discarded without recourse to any statistical tests. If no abnormal conditions can be identified, the discordant value should be reported and indications as to what extent it has been used in the analysis of data should be given.

When an observation that is believed to be an outlier is observed and the physical reasons for its occurrence are unknown, a statistical test may be initiated to lend support to a judgment that a physical reason does actually exist. If on the basis of the test it is concluded that the observation is an outlier, action should be initiated to find the physical reason. The important thing is that the statistical test alone does not tell whether the outlier occurred purely due to chance or due to deviation from prescribed experimental procedure. As such, an outlier should not be rejected on the basis of a statistical test alone.

Figure F-6 shows a flow diagram that summarizes the sequence of logical steps that are helpful to systematically process the treatment of outliers.

In the event a statistical test is required to help identify an outlier, one such test employs Equation (23)

$$T_n = \frac{X_n - \overline{X}}{s}, \qquad (23)$$

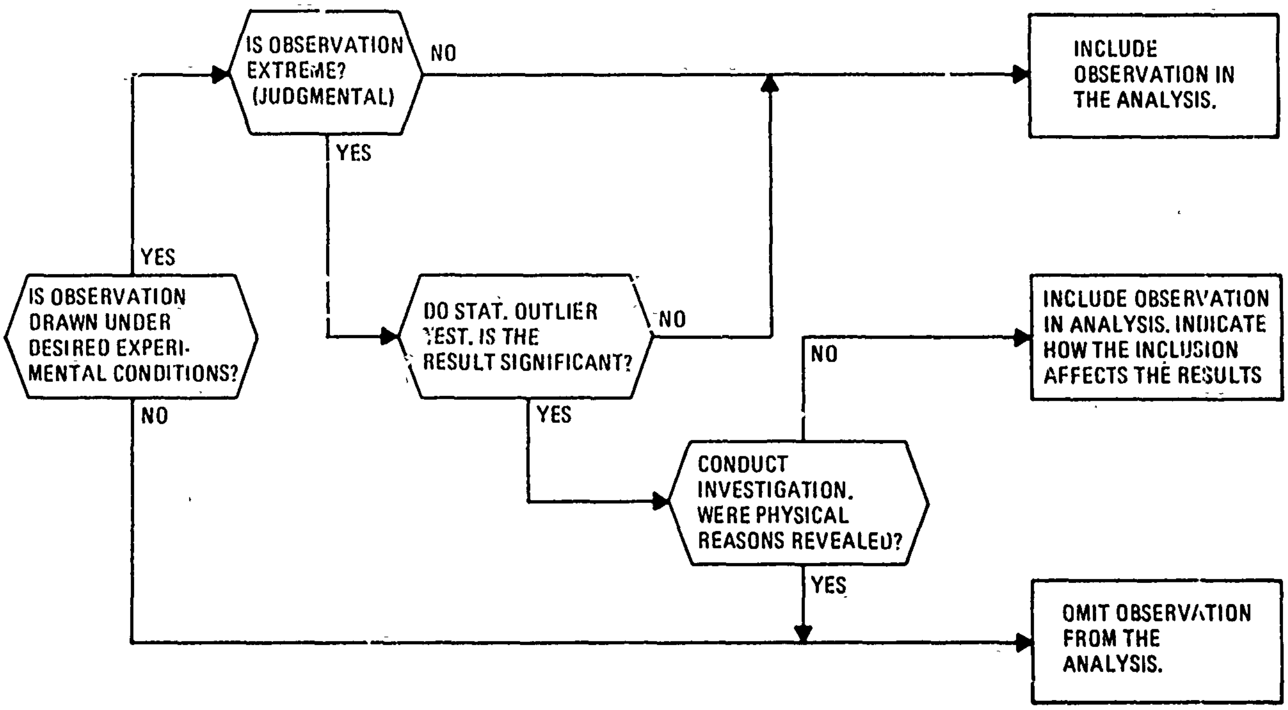


Figure F-6. Flow diagram for the systematic treatment of outliers

where the suspected outlier is the largest-valued observation X_n and $\overline{X}$ and s are the sample mean and standard deviation, respectively. Tables for the values which define the critical region for the test are available in statistical literature.

Transformation of Variables

One of the assumptions accompanying the use of ANOVA models is that the experimental error term be Normally distributed. In the event this assumption is not satisfactorily met, there is the alternative to convert the variable of interest by use of an appropriate transformation function to another variable that may satisfy the Normality requirement.

An example of a transformation of a variable would be an experiment where the "number of successes" out of a given number of attempts would be the dependent variable. Such data is likely to be from a binominal distribution, and the Normality assumption is only approximately met if the average number of successes

is quite high relative to the number of attempts. However, by using "percent successes" as the response variable and, subsequently, transforming it by an arc sine transformation, the Normality assumption may be better satisfied.

Analysis Computer Programs

As part of the data analysis plan, the computer programs to compute the test statistics should be identified. There are numerous software packages available that apply to each of the various types of analysis techniques. For example, a commonly used set of programs are the UCLA (University of California at Los Angeles) Biomedical statistical analysis computer programs. The "Bi-Med Series" includes programs to (1) tabulate and plot test data, (e.g., histogram and x-y plots); (2) perform regression analysis, (e.g., stepwise regression and multiple regression with case combinations); and (3) compute analysis of variance statistics, (e.g., ANOVA for factorial designs and the general linear hypothesis model). The Bi-Med Series has a convenient feature to combine and/or transform (transgeneration) test data variables by a simple designation in the input control data, (e.g., arc sin $\sqrt{x}$ replaces x or x + y replaces x). The transgeneration capability eliminates tedious manual calculations and reduces human errors.

There are computer programs for nonparametric statistical analyses also. For example, for the Sparrow Shoot test program that was conducted through TAWC at Eglin AFB, computer programs were developed for the Fisher exact probability test and the sign test (both with Tocher's modification to test at an exact confidence level).

In summary, there are in existence computer programs that will assist in the analysis of test data; the data analysis plan should reflect the proper interface between the reduced test data and the subsequent analysis with statistical programs.

Interfacing Statistical Results with the Test Evaluation Process

The data anslysis plan should identify how the statistical analysis estimations, comparison, and derived functional relationships relate to the basic test program

objectives and questions. All too frequently, test planners can be so involved with the difficult and challenging details at the technical level that too little thought is given toward the fundamental test objectives.

Appendix G

MATHEMATICS AND STATISTICS

1. CEP

This discussion will be centered on definition of the CEP, proper understanding of that definition (and thereby the usefulness of the CEP), and correct computation or estimation of the CEP under different circumstances. No attempt is made to say whether CEP is the proper measure by which to report test results because to do so would require setting out all the other possible measures of weapons delivery accuracy and describing the advantages and disadvantages of each for different weapon types and different potential data uses.

Definition. The following definition is recommended as a standard for use throughout the Air Force OT&E community:

> CIRCULAR ERROR PROBABLE (CEP). The radius of a circle (centered at the expected arithmetic mean point of impact) within which half the missiles or projectiles are expected to fall.

Explanation of Definition. The CEP is a description of the anticipated distribution of future impacts and thus the radius determined is an estimated value. CEP cannot be the term used for data which only describes the pattern of impacts recorded in a completed test. Furthermore, the term "sample CEP" is a contradiction in terms and should not be used. The term "circular error probable" is actually a more recent way of saying "circular probable error", one of a family of "probable error" parameters. (To conform with the overwhelming trend, the use of "circular probable error" should be avoided). Probable error is defined as the error whose probability of being exceeded is fifty percent. The recommended definition does not say specifically which missiles or projectiles are being described by the estimate or when they will fall, but from the nature of operational testing and the function of statistical estimates, it will be understood that a test is done on a sample of all

missiles plus firing systems of a specific type (and perhaps configuration) and that the Air Force is really interested in what behavior to expect of any missile plus firing system of that type/configuration fired in combat. The launch platform, the physical environment, the attack maneuver, and/or the type of target may also be important. Thus, the missiles or projectiles whose expected behavior is being described by the CEP are from a population characterized by a specific type/configuration and specified combat firing conditions or a specified range of conditions — a population that was sampled to obtain the test data being analyzed. The selection of mean point of impact for the center of the circle simplifies calculations, facilitates understanding of the concept of CEP in the most common statistical terms, and retains significance when bias corrections are made.

Agreement With Established Definitions. The Joint Munitions Effectiveness Manuals (JMEM) define CEP as follows:

> CIRCULAR ERROR PROBABLE (CEP). A measure of delivery accuracy. Its value is equal to the radius of a circle, with its center at the desired mean point of impact, containing half of the impact points of independently aimed bombs, or half of the MPIs resulting from independent aiming operations. (This assumes that the bias is small relative to the CEP in any completely developed system.) The CEP is associated with the circular normal distribution with a standard deviation $\sigma_A = 0.85$ CEP and is a meaningful measure of accuracy if the impact pattern is reasonably circular. As the pattern becomes more elliptical, DEP and REP become the more accurate descriptions of the pattern. (TH 61A1-3-3, p. 1-1)

Although the same page of this manual states that it contains data and equations used in "deriving estimates of combat delivery accuracy, for use in weapon-effectiveness computations," this definition clearly applies to a measure used for describing test data. The JMEM circle is centered on the aim point rather than the mean point of impact; the difference is a single measure of mean deviation from the aim point, or bias, but it is not clear how bias is actually to be dealt with. Using

the recommended definition, bias is specifically removed. While having a single figure to cite for anticipated aim point miss distance has operational appeal, it can be grossly misleading when impacts will cluster several meters (or even tens of meters) from the aim point. The more useful approach would be to report expected bias from the aim point and expected dispersion about the mean point of impact separately. The bias, then, should be corrected once and for all by aiming or hardware changes and the center of the same dispersion pattern would be moved over the aim point. The recommended CEP does not need recalculation for each change in bias. A third notable difference between the recommended and the JMEM definitions is reference to independent aiming operations in the latter. It is true that each impact from salvoed bombs should not be weighted equally with the impact from a bomb dropped alone, but it may be important to know both the dispersion within a salvo of bombs and the dispersion within single bomb passes on a target. Therefore, the recommended definition can be applied to a release of several weapons on independent passes or to the release of several weapons per pass on independent passes, according to the desires of the information-seeker. The JMEM definition of CEP addresses non-circular distribution patterns in a limited way; application of the recommended definition to non-circular patterns will be discussed later.

The Range Commanders Council defines CEP as follows:

> CIRCULAR ERROR PROBABLE (CEP). An indicator of the accuracy of a missile/projectile, used as a factor in determining probable damage to a target; or the radius of a circle within which half the missiles or projectiles are expected to fall. (A Glossary of Range Terminology, document 104-64, p. 23).

This was the model for the recommended definition; it agrees with the recommended definition except for being non-specific about the center of the circle.

The Joint Chiefs of Staff define CEP as follows:

> CIRCULAR ERROR PROBABLE. An indicator of the delivery accuracy of a weapon system, used as a factor in determining

probable damage to a target. It is the radius of a circle within which half of the missiles/projectiles are expected to fall. Also called CEP. (JCS Pub. 1, p. 43).

This definition also agrees with the recommended definition, except for failing to specify the center of the circle.

The Air Force uses the JCS definition of CEP (AFM 11-1, Volume 1, p. 44).

Preliminary Test Data Processing. Several questions must be answered before the proper method for calculation of the CEP can be selected.

(1) Is the population distribution bivariate Normal?

(2) Is the population mean known?

(3) Is the population variance (on each axis) known?

(4) Are the population variances equal?

Standard methods for making these decisions are prescribed below.

Population Mean. By stating that the population μ is known, the analyst is asserting that he either believes it to be the aim point or believes he knows the bias exactly. The latter case is not likely because a known bias probably would be corrected for. Initially, then, the analyst assumes that the population mean is at the aim point and that any bias apparent from the test data is only due to the sampling process.

Hotelling's Generalized T^2 statistic, given by

$$T^2 = \frac{S_y^2 N \bar{x}^2 - S_{xy} N \overline{xy} + S_x^2 N \bar{y}^2}{S_x^2 S_y^2 - S_{xy}^2}$$

can be applied to tell whether it is reasonable to accept this hypothesis of zero bias. A comparison of the quantity

$$F = \frac{(N-2)\, T^2}{2(N-1)}$$

with $F_{.05}$ (2, N-2) as tabulated will reject the hypothesis only when the bias is quite surely non-zero.

If the non-zero bias hypothesis must be rejected, the analyst can no longer state that the population mean is known; although the sample mean is an unbiased estimate of the population mean, the only information available is from that one sample. The proper procedure now is to report the apparent bias and calculate variances from the sample mean.

Population Distribution. Most statistical techniques applicable to this problem assume that the variations along the two axes are uncorrelated. Thus, the initial hypothesis is that the data do not provide reason to doubt zero correlation. If it can be shown that there is reasonable likelihood that the data is from a bivariate Normal distribution (by means of an X^2 test for goodness of fit, with .05 significance level) then the sample data can be tested for correlation with the statistic

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

with N-2 degrees of freedom for sample size N and sample correlation coefficient r. Again, a significance level of .05 should be used to assure that the zero correlation coefficient (population parameter) hypothesis is not rejected without convincing reason. If correlation does exist, it can be removed (for calculation purposes) by rotation of the axis through an angle of

$$\theta = 1/2 \tan^{-1} \frac{2rs_xs_y}{s_x^2 - s_y^2};$$

in other words by calculating new x' and y' values

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

When the data is not believed to be Normal, Pitman's Test may be used to check for correlation. Even in the distribution-free case, the technique for removing correlation is the same.

Population Variance. The initial hypothesis should be of some particular value for the variance along each axis, based on system design information. When variances are not shown to be different from this hypothesis on the basis of the OT&E results (another F-test), it will be stated that the variance (σ^2) is known. If the initial hypothesis does not hold up under test, the population variance must be estimated from the test statistic s^2. One should be cautious about acceptance of a system design quote of variance; however, when it is based on DT&E results rather than successful system design, it may be no more valid as a population parameter than the OT&E estimate. When correlation has been removed by rotation of the data to new axes, the variance cannot be known, because all the information available after this transformation is that from the test results. These tests should also inform a person whether or not the distribution is circular by identifying agreement with or statistically - significant differences from the variances initially presumed (if any).

Calculation of CEP — Circular Normal Data. Four cases are possible: (a) known μ and σ; (b) known μ, unknown σ; (c) unknown μ, known σ; (d) unknown μ and σ.

a. CASE (a). The first case is straightforward. When both of the population parameters are known, the radius of the circle containing 50% of the population is exactly

$$CEP = [\ell n\ 4]^{1/2}\sigma .$$

It is worth mentioning that in this case the test data provides no new information, except by way of confirmation of the initial hypothesis, and thus this CEP calculation (not an estimate at all) could be made from data available before the test.

b. CASE (b). When σ alone is unknown, the estimate of CEP is

$$CEP_{est} = \frac{[\ell n\ 2]^{1/2}\ \Gamma(N)}{\Gamma(N + 1/2)} \left[\Sigma(X-\mu_x)^2 + \Sigma(Y-\mu_y)^2\right]^{1/2}$$

N is the number of impact points in the sample and $\Gamma(Z)$ is the complete gamma function. This is the most efficient (i.e., minimum variance) unbiased estimate possible.

c. CASE (c). When μ is unknown, it is probably going to be assumed that σ is unknown. Th problem with σ known and μ unknown is not inconceivable--only difficult to rationalize. When it is desired to estimate the CEP on this basis, the formula is the same as for CASE a.

d. CASE (d). In the unknown μ, unknown σ case, the unbiased estimate is

$$CEP_{est} = \frac{[(N-1)\ \ell n\ 2]^{1/2}\ \Gamma(N-1)}{\Gamma(N-1/2)} \left[s_x^2 + s_y^2\right]^{1/2}$$

where S_x^2 is the sample estimate of the population variance

$$s_x^2 = \frac{\sum_{i=1}^{N}(x-\bar{x})^2}{N-1}$$

Calculation of CEP — Elliptical Normal Data. As above, four cases are possible.

a. CASE (a). μ, σ_x and σ_y known. Tables can be used to look up the exact CEP. The most extensive of these is published as "A Method for Computing the Generalized Circular Error Function and the Circular Coverage Function" (A.R. Di Donato and M.P. Jarnigan, U.S. Naval Weapons Laboratory Report No. 1768, 23 January 1968.) The table in Appendix C should be entered with V = 0.50.

b. CASE (b). When μ is known but σ_x and σ_y must be estimated, the minimum-variance unbiased estimate is

$$CEP_{est} = \frac{[\ell n\ 4]^{1/2}\ \Gamma(N)}{2\,\Gamma(N+1/2)} \left\{\left[\sum (x-\mu_x)^2\right]^{1/2} + \left[\sum (y-\mu_y)^2\right]^{1/2}\right\}$$

c. CASE (c). An estimate of the same parameter for the σ known, μ unknown case is the same as for CASE (a).

d. CASE (d). When σ_x, σ_y and μ are unknown, the unbiased estimate of CEP is

$$CEP_{est} = \frac{[(N-1) \ln 4]^{1/2} \Gamma(N-1)}{2 \Gamma(N - 1/2)} \left[S_x + S_y\right]$$

Comparison With Other Methods Of Calculation — Tolerance Regions. A tolerance region also contains a certain percent of the population values (50% if desired) but a 50% tolerance region is calculated in a different manner from the CEP equations above and it has (the non-statistician would say) a subtly different meaning. The concept is attractive because the same formulas can be used to calculate 50% containment ellipses for non-circular distributions (although not CEP's for non-circular distributions), and the same type of estimation can be made for a non-normal population. It is extremely unlikely that anyone using the CEP as a measure of weapons delivery accuracy could make and support a choice between the two approaches on the basis of the mathematical differences, so the recommended standard formulas were chosen in recognition of the fact that they are most like many of the formulas that are in current use (whether used properly or not) and therefore they would be accepted most readily for general use. The CEP's calculated by the two methods converge for large samples and are different as shown in Figure G-1 for small samples. Figure G-1 shows only the CASE (d) comparison. Results for CASES (b) and (c) are similar; the CASE (a) solution is exact, not an estimation, and the results are identical. The reason these two approaches can be different is that they are both estimates of a value based on incomplete information and they make different assumptions about the correct way to predict the generally accurate result.

DEP and REP — Univariate Measures. At times, one-dimensional 50% containment intervals are used because of preferences in the case of highly elliptical impact patterns

or because of direct applicability in the case of rectangular targets. Definitions analogous to the recommended CEP definition are:

RANGE ERROR PROBABLE (REP). Half the distance between two imaginary lines on the ground, perpendicular to the aircraft approach line and equidistant from the expected arithmetic mean point of impact, between which half the missiles or projectiles are expected to fall.

DEFLECTION ERROR PROBABLE (DEP). Half the distance between two imaginary lines on the ground, parallel to the aircraft approach line and equidistant from the expected arithmetic mean point of impact, between which half the missiles or projectiles are expected to fall.

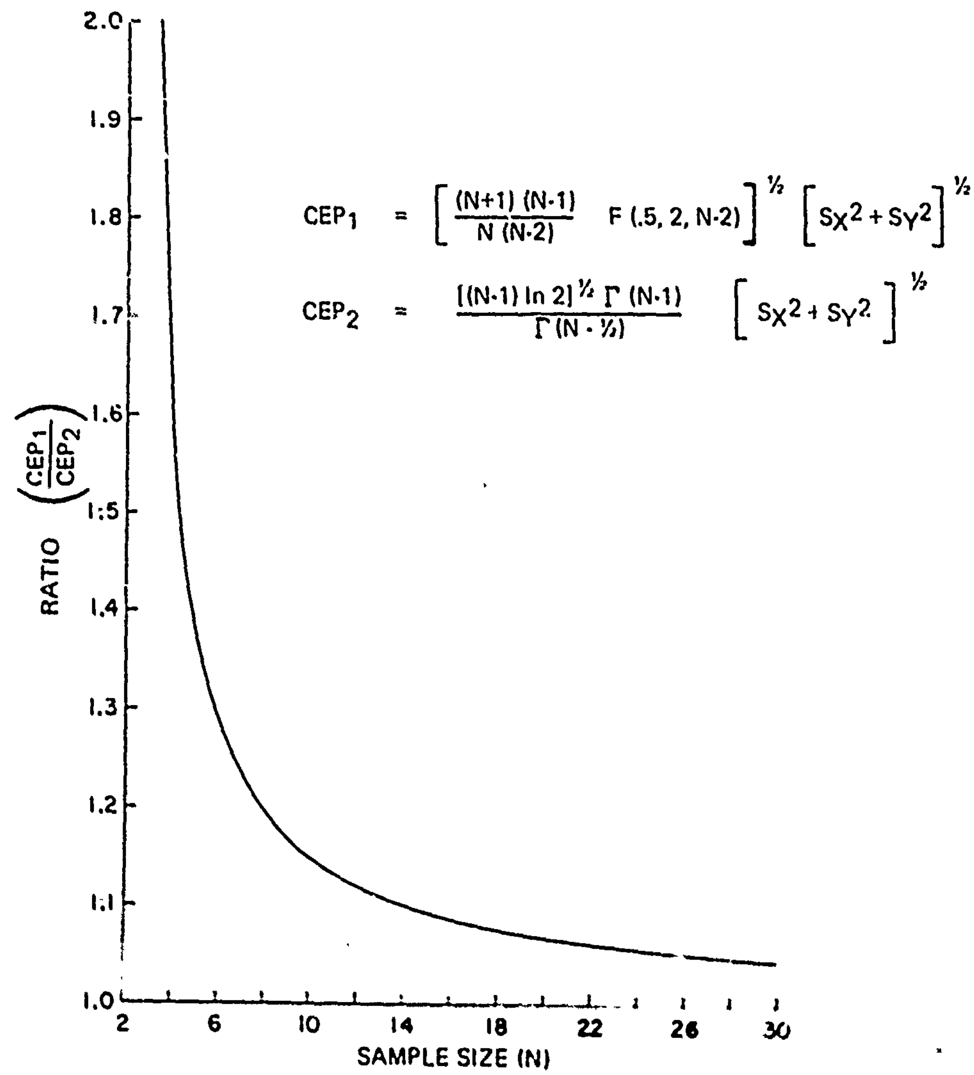


Figure G-1. Comparison of CEP estimates (μ unknown σ unknown)

Note that the rectangle formed by overlaying the boundaries for the DEP band and the REP band contains 25% of the population (50% x 50%), not 50%. The exact formulas (μ and σ known) for the defined terms are:

$$\text{REP} \simeq 0.6745\ \sigma_R$$

$$\text{DEP} \simeq 0.6745\ \sigma_D$$

but the equal probability bands that would contain a 50% rectangle are given by

$$A \simeq 1.052\ \sigma_R$$

$$B \simeq 1.052\ \sigma_D$$

<u>Interval Estimate.</u> Confidence limits may be placed on the point estimate of CEP (thus defining a minimum and maximum area circle for containment of 50% of the population) in the same way, they are put on point estimates of the standard deviation. The desired confidence level is chosen, the number of degrees of freedom is calculated (exactly or approximately) from a formula in Figure G-2 and the appropriate interval limit is computed by the relation

$$\text{Interval limit} = \text{CEP}_{\text{point est}} \times \left[\frac{\nu}{\chi^2\,(\text{confidence level}, \nu)}\right]^{1/2}$$

ν is the number of degrees of freedom. There is of course no interval estimate for CASE 1 (μ and σ known) since the distribution is assumed to be known exactly.

<u>Reporting The CEP.</u> To provide complete information of weapons delivery accuracy, several items of information should be reported along with the CEP itself. The following information should be given:

(1) Assumptions regarding prior data and the bias, circularity, and Normality of the test data parent population.

(2) Method used for calculating exact CEP or estimate.

(3) Sample size.

Assumed Distribution	Case	
Circular	2	2N
	3	2N-2
	4	2N-2
Elliptical	2	$\left(1 + \frac{s^2 \text{ small}}{s^2 \text{ large}}\right) N$
	3	$\left(1 + \frac{s^2 \text{ small}}{s^2 \text{ large}}\right) (N-1)$
	4	$\left(1 + \frac{s^2 \text{ small}}{s^2 \text{ large}}\right) (N-1)$

Figure G-2. Degrees of freedom (ν) for interval estimates of CEP

(4) Conditions under which data obtained, including altitude at release.

(5) Variance along each axis or the estimates thereof.

(6) CEP (exact or point and interval estimates).

Step-By-Step Procedure For Calculation Of CEP

1. Calculate range and crossrange deviations from aim point for each impact.

2. Use Hotellings Generalized T^2 statistic associated with an F test (.05 significance level) to see if any apparent non-zero bias is statistically significant. Use aim point for population mean if it is not; use mean point of impact for estimate of population mean if it is.

3. Calculate the variance point estimate on each axis if the population mean is estimated.

4. Check the data for statistically - significant non-Normality with a χ^2 test (.05 significance level).

5. Check for statistically significant correlation with a t test (Normal data) or the appropriate nonparametric test (.05 significance level). Rotate axes to remove statistically significant correlation for calculation purposes.

6. Check for statistically - significant differences from the assumed values of variance on each axis with the F test (.05 significance level).

7. Choose appropriate calculation method.

8. Calculate the CEP.

2. CONFIDENCE LIMITS

Whenever parameters of a population (such as mean miss-distance of all AGM-74 missiles fired at tanks from the F-18 while below 1000 meters and within 2 km slant range of the target) are measured by conducting a test on a sample (say 10 missile firings) from that population, one cannot be sure that the mean value calculated from the sample will be exactly the same as the mean for all missiles that would ever be fired. It might be expected, however, that the sample mean could be used as an estimate of the population mean if the sample were obtained by random selection from all AGM-74 missiles and all firing scenarios within the bounds imposed. It might further be expected that as more missiles were fired (the sample became larger), more confidence could be placed in the closeness of the estimate to the true population value. The test officer realizes that there is a distribution of miss distances due to the myriad of variations in missile construction and circumstances under which the missile is fired, but the more missiles he fires and the greater the range of scenario variables under which they are fired (within the understood population, of course) the more he will feel that he has sampled the population well and observed something representative of the entire population rather than the chance occurrence of (say) the two longest miss distances that would ever be realized.

There is a way to quantify the intuitive feelings of this officer and tell just what the possibility is that the sample mean was derived from the chance occurrence of a group of miss distances all from the high end of the distribution. A confidence interval can be constructed about the sample mean and the statement made that with a level of confidence of 95% (for example) the true population mean lies within that interval. This says that there is only one chance in twenty that the population mean lies outside the quoted interval. Now, in fact the population mean either does or does not lie inside the quoted interval, but the interval is being constructed on the basis of incomplete information (a sample) and the confidence level gives the probability of the incomplete information providing the right result. The statistical

calculation will tell what the probability is that the obtained value represents nothing more than chance observations all tending toward the high (or low) end of the distribution. Four things interact in these calculations: the variation inherent in the population distribution, the size of the sample taken, the length of the confidence interval, and the level of confidence. Consider the following example:

Example. Twenty missiles are fired at similar targets in similar scenarios. The recorded miss-distances are shown in Figure G-3. The arithmetic mean miss distance for this sample is 7.0 meters. It can be stated now with 95% confidence that the population mean (the true average miss distance) is between 4.5 and 9.5 meters.

4.2	5.4	13.0	1.2
6.0	6.6	8.4	13.8
2.4	3.4	5.8	1.9
12.2	0.1	10.5	7.8
12.9	1.1	20.4	2.9

Figure G-3. Miss distance (meters)

This statement informs the Test Officer that there is still one chance in 20 that the miss distances observed come from a population with a mean miss distance less than 4.5 meters or greater than 9.5 meters. This is a symmetrical two-sided confidence interval; i.e. coefficients to allow construction of these intervals are tabulated. By moving both confidence limits in the same direction (increase both or decrease both), although not by equal amounts, additional (longer) 95% confidence intervals can be constructed.

Suppose, however, that the Test Officer is only interested in the upper limit on mean miss distance; then a one-sided confidence interval can be constructed. In the present example a one-sided 95% confidence interval includes all values up to 9.1 meters: there is one chance in 20 that the true population mean is greater than 9.1 meters. Coefficients for other confidence levels are also tabulated. The one-sided 90% confidence interval goes to 8.6 and the one-sided 99% confidence interval goes to 10.0.

How much better off would the Test Officer be if he could fire more missiles? Assuming that the sample mean miss-distance and the sample estimate of the population variation (the sample estimate of standard deviation is what is actually used) are unchanged in a sample of 121 firings, it could be stated with 95% confidence that the miss distance is less than 7.8 meters. The improvement in quality of the estimate may or may not be worth the expenditure in resources to fire an additional 101 missiles. The change in size of the confidence interval is not linear with sample size, however. For if only 5 missiles had been fired (assuming "similar" test results), the one-sided 95% confidence limit would be 12.1 meters, and if one-sided, 95% confidence limit would be 30.9 meters!

It may be that the Air Force is not so interested in the population mean miss distance as in the probability that a certain percentage of all the firings in the population will have a maximum miss-distance of a certain value. This value would be a tolerance limit and it would bound a tolerance interval. While bearing a resemblance to confidence limits and confidence intervals, they are not the same. In the same example*, the one-sided 95% tolerance limit for 90% of all firings is 19.4 meters. The Test Officer is assured that there is only one chance in 20 of being in error if he says that 90% of all AGM-74 missiles fired under these conditions will miss by no more than 17.3 meters. Two-sided tolerance intervals can also be constructed. Tolerance interval coefficients are tabulated in much the way that confidence interval coefficients are.

A third type of probability interval is the prediction interval. This interval has a stated probability of containing the miss distance (in the current example) of the next single test. In this case, the one-sided 95% prediction interval for the next (twenty-first) miss distance is bounded by 16.5 meters. That is, there is only one chance in 20 that these results are misleading and that the next missile will be off target by more than 16.5 meters.

*Computation of tolerance intervals requires some statement about the form of the distribution of values in the population. To continue the present example a Normal distribution is assumed.

Figure G-4 gives a comparison of the one-sided 95% probability interval limits for the example discussed herein:

confidence limit	9.1
prediction limit	16.5
tolerance limit (90% pop.)	17.3

Figure G-4. One-sided 95% probability interval limits (meters).

Discussion to this point has been devoted to probability intervals for arithmetic means. Confidence intervals can also be constructed for the median (50th percentile) and other quantiles, for proportions and percentages, for the statistical range, and for the standard deviation as well as other attributes that can be grouped, ranked, scaled, or measured. Joint confidence regions for the mean and standard deviation (allowing a single probability statement concerning both) can be constructed. In contrast to the situation for the arithmetic mean, the shortest confidence interval for the standard deviation will not be symmetric in location about the value on which it is based. Thus, although the probability that the population standard deviation is less than the lower interval limit equals the probability that it exceeds the upper interval limit, the two limits are not equidistant from the observed estimate of the standard deviation.

Multi-dimensional probability regions can be calculated for simple region shapes (circle, ellipse, rectangle).

The calculation of probability regions is relatively straightforward just as long as the assumptions incorporated as to the shape of the distribution and the completeness of knowledge about the mathematical description of that distribution are kept track of and communicated.

One problem may arise in connection with that completeness of knowledge, however. If it can be assumed that the true population standard deviation is known, a different statistic may be used in the calculation of intervals or observed significance levels, giving shorter intervals and lower significance levels. This is fine if the sample

size is large enough for the confidence level at which one is working; the often-quoted sample size of 30 for which it is acceptable to use the z statistic instead of Student's t statistic introduces errors of 2.3% in the length of the confidence interval at 90% confidence levels and 6.5% at 99.5% confidence levels. With the availability of calculating machines and tabulations of Student's t values for larger sample sizes, there is no excuse for not using Student's t in calculations involving samples of size 150 and greater to keep the error in making statistical statements as low as possible. Figure G-5 shows graphically the error in confidence interval length for a range of sample sizes and several confidence levels if z is used when t should be.

It is sometimes difficult to decide what confidence level to use. Of course, the higher the confidence level the smaller the chance that the calculated confidence interval does not contain the population parameter being estimated, but to gain increased confidence in an interval of any given length the sample size, (i.e., cost of the test) must be increased. An alternative solution is to back off on the interval size requirement. After all, is it operationally important to know the mean miss distance within 4 meters, or would 6 meters be an adequate estimate? Or is it operationally important to have 99% confidence in the answer presented rather than only 95% or 90%? Some compromises can be made. Then if it cannot be decided whether it would be more helpful to report the longer, higher-confidence interval or the shorter, lower-confidence level, it would be wise to show a "picture" of the data as in Figure G-6. This shows the confidence limit for any confidence level between 50% and 99.9% in one graph. No additional data is required -- just a few calculations to show the same data in different ways.

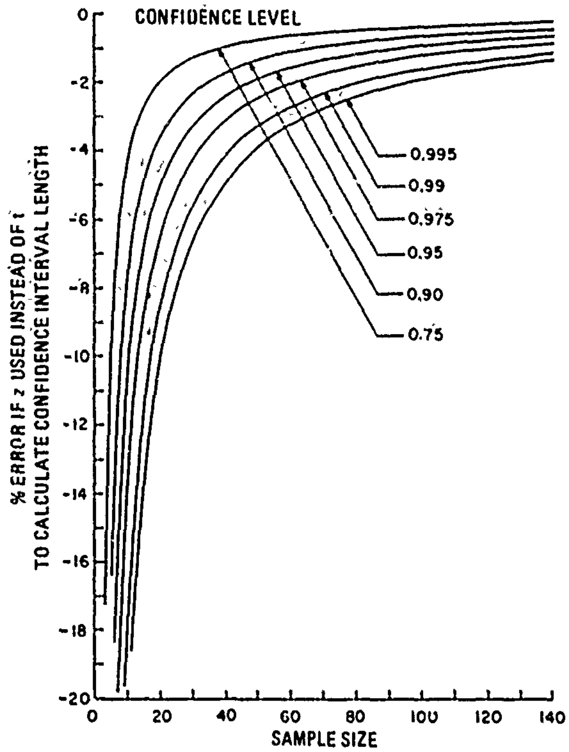


Figure G-5. Percent error if z is used instead of t

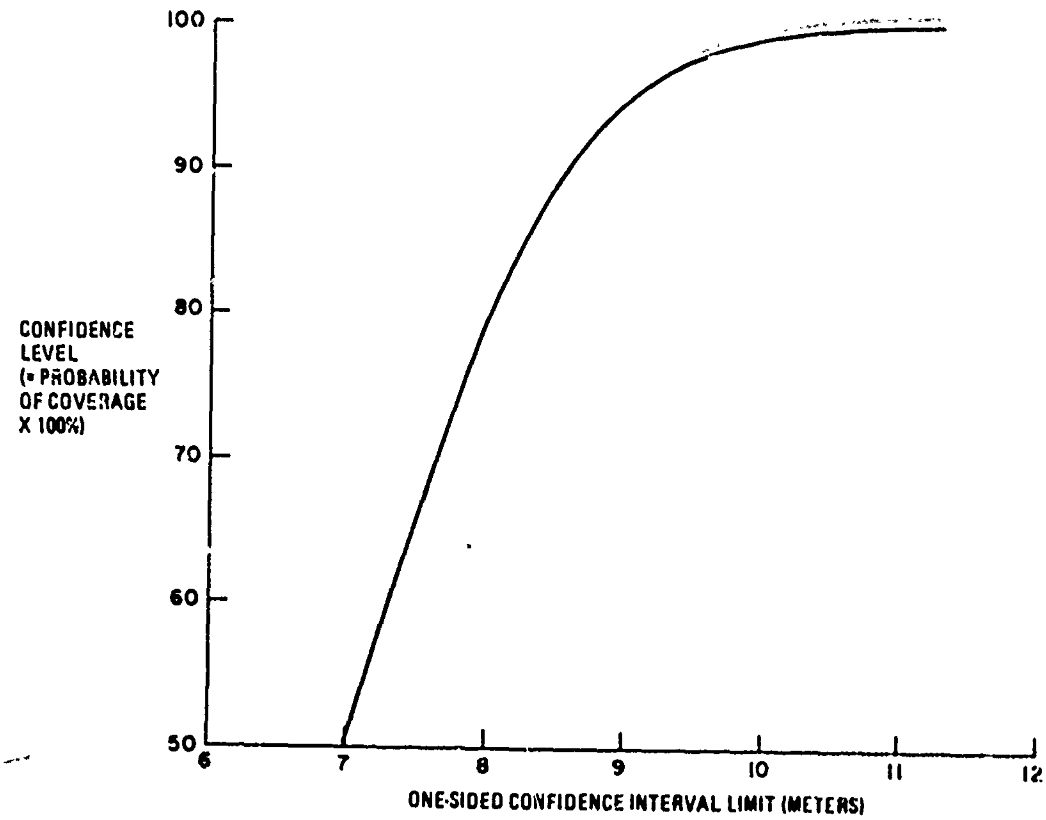


Figure G-6. Graphic display of interval estimate for mean miss distance from Figure G-3 data

3. SIGNAL-TO-NOISE

The signal-to-noise ratio is one of the most pervasive of all engineering parameters. It is used to describe conditions in optics, radio transmission, speech intelligibility, and radar systems, among others. Yet in spite of the widespread acceptance and use of the term (in fact, probably because of it) there is often confusion about the true parameter that is being discussed.

Signal-to-noise, usually written in symbols as S/N, is the ratio of signal power to noise power at a selected point in the system. There is no unique signal-to-noise ratio which refers to an entire system; any source of degradation (i.e., increased noise) changes the signal-to-noise ratio, so a system may be described by as many different signal-to-noise ratios as there are nodes through which signals pass.

If signal power and noise power were constants at every node, signal-to-noise would be the simple ratio of two constants. Both signal power and noise power vary in time, however, and a standard method of measuring each must be recognized. It is standard to use the peak power of a periodically varying signal and the root-mean-square noise power of random noise or the peak value of impulse noise. Where there is a possibility of ambiguity, or if another measure of signal and noise is needed, definition of the signal and the noise must be associated with the term (e.g., peak signal-to peak noise). Signal-to-noise for an optical system is the ratio

$$\frac{\text{intensity of signal}}{\text{total intensity of all other contributors to intensity}}$$

4. JAMMING POWER

Jamming power can be defined simply as a measurement of the electro-magnetic energy radiated from a jamming system per unit time. While this quantity may be of fundamental importance to the electrical engineer developing a jammer, the fundamental interest of the operational tester is in whether the jammer effectively degrades the enemy's ability to communicate, to guide his missiles, to detect and track aircraft and missiles, to navigate, etc.

If a jamming system has the (loosely-defined) power or simply ability to effectively jam a particular enemy system, the operator and the operational tester should not be concerned with measurement of the radiated electromagnetic energy. The operator's jamming power is measured by the ability to cause degradation in some data element (or elements) of the Measure of Effectiveness applicable to the system being jammed.

[This should not prohibit the operational tester for measuring radiated electromagnetic energy as a diagnostic device if he so desires.]

5. INTELLIGIBILITY

Intelligibility vs Other Measures

Intelligibility, although not the only measure of speech quality, is the "sine qua non" of speech communication systems. Other measures, such as loudness, naturalness and speaker identifiability, may be important to the operational effectiveness and suitability evaluation of a particular system, but it is the ability of the system to convey information that most directly reflects a speech-communication system's purpose. There may or may not be an overlap between intelligibility and these other factors for a particular speech-communication system, depending on its function.

Intelligibility vs Articulation

When used as a measured property of a system, not colloquially, intelligibility is the percent of meaningful words, phrases or sentences spoken by a talker (or talkers) that are heard correctly by a listener (or listeners). It is distinguished from articulation, which is used when the units of speech material are meaningless syllables or fragments. While very useful tests of speech-communication systems based on physical measurements alone have been developed, an operational test of a speech-communication system must involve live talkers and listeners. (This does not completely preclude the use of recorded speech, but forces limits on the use of recordings.) The articulation index, a physical measure based on the intensity of sound in each of 20 equally-important speech frequency bands, cannot be relied upon in operational testing any more than wind tunnel tests of aircraft performance. In spite of the success of the articulation index, it measures response to a single composite speech spectrum with the microphone held a specified distance from the lips and in general "averages" speech-communication scenarios into one artificial test.

Talker-Link-Listener Systems

Depending on the objectives of a given operational test, intelligibility can be used in an evaluation of a link between talker and listener or of a talker-link-listener system. In one case, variability in talkers and listeners would be so well characterized that it

would not have to be included as a part of the test, but in the other it would be vitally important to include a variety of talkers and listeners. More will be said on this subject later; the intent here is to point out the importance of the talker-link-listener system concept.

Choice of Test

The function of a particular speech-communication system should be carefully considered in selection of the proper test of effectiveness and suitability. The material to be communicated in operational employment should be the basis for the test. If messages over the system consist of single sentences, single sentences should be used in the test. If they consist only of short phrases or single words, short phrases or single words should be used. If a specialized language will be used (e.g., air traffic control), messages in that language should be used. These considerations should not be implemented as restrictions on testing of the system at all because if a speech-communication system needs great flexibility and versatility it should be tested in a number of different uses. The number of variables in speech communication and in speech communication testing is impressive, however, and in overlooking the complexity of the subject a tester may overlook the need to tailor a test to the expected operational employment of a system. The following is a list of variables in speech-communication intelligibility which should be acknowledged in the design of a test; it is probably not complete.

1. Message Length

The meaning of a message is more apt to be clear if the context can be used to identify missing words than if each word stands in isolation.

2. Word Category

When words are predictable because they are known in advance to be from a certain set (e.g., part numbers, street addresses), intelligibility is increased.

3. Word Familiarity

Words that are foreign to either the talker or the listener are less intelligible.

4. Word Length

Words that are several syllables in length put each syllable in a context of its own. The decreasing number of longer words and the unique rhythm of long words make them more intelligible.

5. Talker position with respect to microphone.

Microphone efficiency may vary as the origin of the sound moves in an arc a uniform distance away.

6. Talker Orientation with Respect to Microphone

A talker who speaks directly into a microphone will be more easily understood than one who faces away from it. Reflections off walls or other physical surfaces may degrade part of the frequency spectrum as well as cause echoes.

7. Talker Stress

Distractions due to other concerns cause the talker to lose concentration on the necessary pronunciation and enunciation of each word in a message.

8. Talker Background Noise

Other sounds received may compete for the listener's attention as well as mask the message he is supposed to hear.

9. Syllable Length

Syllables that are stressed and lengthened come across more clearly.

10. Talker Dialect or Accent

Certain talkers are not as universally easy to understand.

11. Other talker voice qualities.

Pitch affects the intelligibility of a message and this is part of the reason that male voices are in general preferred over female voices in the talker role. Other talker-talker differences are known to exist, but have been harder to identify a priori.

12. Talker Training

Training in the use of speech communication systems increases intelligibility by lessening the differences due to some of the effects discussed above. Even informal training, such as that gained in the repetition of an intelligibility test, plays a role.

13. Listener Position with Respect to Speakers

There are fewer losses and echoes if the listener is wearing earphones than if he is listening to sound coming from a speaker across a room.

14. Listener Orientation with Respect to Speaker

Like microphones, speakers are not equally efficient in all directions. The use of multiple speakers for different communication systems may be confusing to the listener.

15. Listener Stress

16. Listener Background Noise

The tester must decide how the operational employment of a particular speech-communication system involves each of these variables. Then there are other sources of variability in the test results due solely to the method of test itself.

1. Listener Motivation.

If a test becomes too routine, the listener will lose interest and perform poorly.

2. Listener Memorization

A good sample of talker and message combinations could be recorded and used repeatedly. Each listener's response must be spontaneous, however, and the tester must ever guard against the possibility that listeners will become so familiar with the messages being used that the response is based on that familiarity rather than on the intelligibility in a particular trial. The use of different talkers seems to hinder memorization of a particular message.

Reporting Intelligibility Test Results

Many aspects of the intelligibility test must be reported in order to convey the full significance of the results. In addition to a statement of the intelligibility realized (where an interval estimate of the percent is necessary to show the degree of uncertainty) the scope of the system considered a part of test and the handling accorded different potential sources of variability in the system should be reported. The satistical analysis of data should be reported also, to the extent that it reveals sources of statistically significant variability in the system performance and to the extent that it reveals the source of system deficiencies. The method of control over the test, such as comparison of the intelligibility of recordings made at the talker with the intelligibility of the same trial-message at the listener end of the system should be reported.

6. STATISTICAL INDEPENDENCE

Two events are said to be statistically independent if

$$P(AB) = P(A)P(B)$$

where P(A) is the probability of event A occurring, P(B) is the probability of event B occurring, and P(AB) is the probability of events A and B both occurring. For example, suppose we have a standard deck of cards and draw one card at random and then draw another. What is the probability of drawing one spade and one club? Let event A be drawing a spade and event B be drawing a club. Since the events are independent, the order can be (1) draw a spade then a club, or (2) a club and then a spade. Thus

$P(AB) = \frac{13}{52} \cdot \frac{13}{51} + \frac{13}{52} \cdot \frac{13}{51}$ where the first represents drawing a spade then a club and the second, a club then a spade. Or

$$P(AB) = .12745$$

The above example also illustrates the mutual exclusiveness of the two ORDERED events A' and B' where A' is defined to be the ordered event of drawing a spade and then a club, and B' is defined to be the ordered event of drawing a club and then a spade. With these definitions, $P(A' + B') = P(A') + P(B')$. This defines the probability of at least one ordered event occurring. Therefore, $P(AB) = P(A' + B') = P(A)\,P(B)$ for the case illustrated.

7. CONDITIONAL PROBABILITIES

Let $P(Y|X)$ mean the probability that event Y occurs, given the fact that event X occurred. $P(X)$ means the probability that event X occurred. Thus Y is conditional upon X occurring. Then the probability of X and Y both occurring, when Y is condition upon X [represented by $P(XY)$] is

$$P(XY) = P(X)P(Y|X).$$

For example, suppose we have a standard deck of cards and draw one card at random and then another. What is the probability of drawing a spade (event X) and then a club (event Y)?

$$P(X) = \frac{13}{52} \text{ and } P(Y|X) = \frac{13}{51}$$

Thus the probability of X and Y, where Y is conditional upon X is

$$P(XY) = \frac{13}{52} \cdot \frac{13}{51} = .06373$$

Notice that the result is quite different than calculating the probability where the events were independent.

Appendix H

GLOSSARY

1. LIST OF ACRONYMS

A - Availability

ACB - Budget Directorate

ACP - Area Coordinating Paper

ACQ - Acquisition

ADC - Aerospace Defense Command

ADCP - Air Defense Command Post

ADP - Automatic Data Processing

ADPE - Automatic Data Processing Equipment

ADPS - Automatic Data Processing System

ADS - Automatic Data System

ADTC - Armament Development and Test Center

AF - Air Force

AFAG - Air Force Advisory Group

AFFTC - Air Force Flight Test Center

AFLC - Air Force Logistics Command

AFOT&E - Air Force Operational Test and Evaluation

AFPE - Air Force Preliminary Evaluation

AFR - Air Force Regulation

AFSC - Air Force Systems Command

AFSCC - Air Force Special Communications Center

AFSS - Air Force Security Service

AFTEC - Air Force Test and Evaluation Center

AFTEC/cc - The Commanding General, AFTEC

AGE - Aerospace Ground Equipment

ALC - Air Logistic Center

ASD - Aeronautical System Division of AFSC

ASIP - Aircraft Structural Integrity Program

ATC - Air Training Command

Az - El - Azimuth-Elevation

LIST OF ACRONYMS (Cont)

BA - Budget Authorization

BCI - Budgetary Cost Information

BPE - Best Preliminary Estimate

C - Capability

C^3 - Command, Control, and Communications

CBU - Cluster Bomb Unit

CDR - Critical Design Review

CE - Communications/Electronics

CEP - Circular Error Probable

CI - Configuration Item

CIDT&E - Combined IOT&E and DT&E Program

CM - Countermeasures

COMSEC - Communications Security

D - Dependability

DCP - Development Concept Paper

DCS - Deputy Chief of Staff

DDC - Defense Documentation Center

DDOI - Deputy Director Operating Instruction

DDR&E - Director of Defense Research and Engineering

DIFM - Due-in from Maintenance

DMS - Data Management System

DMIS - Data Management Information System

DOD - Department of Defense

DPD - Data Project Directive

DPP - Data Project Plan

DSARC - Defense Systems Acquisition Review Council

DT&E - Development Test and Evaluation

DTL - Data Profile Time Line Chart

ECM - Electronic Countermeasures

EDP - Electronic Data Processing

EMI - Electromagnetic Interference

EOD - Explosive Ordnance Disposal

EW - Electronic Warfare

FAA - Federal Aviation Agency

FDR - Flight Data Recorder

FOT&E - Follow-on OT&E

FTC - Flight Test Center

FYDP - Five Year Defense Program

F&FP - Force and Financial Program

GFE - Government Furnished Equipment

HOI - Headquarters Operating Instruction

HQ USAF - Headquarters, United States Air Force

I&L - Assistance Secretary of Defense (ASD), Installation and Logistics

ILSP - Integrated Logistics Support Plan

IOC - Initial Operational Capability

IOT&E - Initial OT&E

JOT&E - Joint Operational Test and Evaluation

JOTR - Joint Operational and Technical Review

JTD - Joint Test Director

LCC - Life Cycle Costs

LGM - Director of Maintenance Engineering

MAC - Military Aircraft Command

MAJCOM - Major Command

MASF - Military Assistance Service Funded

MDC - Maintenance Data Collection

LIST OF ACRONYMS (Cont)

MIL STD - Military Standard

MIS - Management Information System

MISEDS - Machine Independent Systems Effectiveness Data System

MOD - Modification

MOE - Measure of Effectiveness

MTBF - Mean Time Between Failure

MTTF - Mean Time to Failure

NDI - Non-Destruct Inspection

NORS - Not Operationally Ready Supply Support

NTIS - National Technical Information Service

O&M - Operations and Maintenance

OD - Operation Directive

OPR - Office of Primary Responsibility

OR - Operations Requirement

OSD - Office of Secretary of Defense

OT&E - Operational Test and Evaluation

PA - Program Authorization

PAD - Program Action Directive

PCA - Physical Configuration Audit

PDM - Program Decision Memorandum

PDR - Preliminary Design Review

PEM - Program Element Monitor

PID - Program Introduction Document

PM - Program Memorandum

PMD - Program Management Directive

PMP - Program Management Plan

PMI - Preventive Maintenance Instruction

PO - Program Office

LIST OF ACRONYMS (Cont)

POL - Petroleum, Oils, Lubricants

POM - Program Objective Memorandum

PPBS - Planning, Programming and Budgeting System

PRD - Program Requirements Document

PSP - Program Support Plan

QRC - Quick Reaction Capability

RCC - Range Commander's Council

R&D - Research and Development

RDG - Assistant for RDA Programming

RDP - Director of Development and Acquisition

RDQLM - Requirements Review Group Secretariat

RDR - Directorate of Reconnaissance and Electronic Warfare

RFP - Request for Proposal

R&M - Reliability and Maintainability

ROC - Required Operational Capability

RPV - Remotely Piloted Vehicle

RRG - Requirements Review Group

SA - Studies and Analysis

SAC - Strategic Air Command

SAF - Secretary of the Air Force

SAM - Surface-to-Air Missile

SAR - Selected Acquisition Report

SECDEF - Secretary of Defense

SM - System Manager

SMD - System Management Directive

SOD - Support Officer for Data

SOR - Specific Operational Requirement

SOW - Statement of Work

SPO - System Program Office

SPP - System Package Program

SPR - System Program Review

SSA - Source Selection Authority

SSAC - Source Selection Advisory Committee

SSB - Source Selection Board

STINFOR - Scientific and Technical Information

TAC - Tactical Air Command

TAWC - Tactical Air Warfare Center

T&E - Test and Evaluation

TEFB - Test and Evaluation Facilities Base

TFWC - Tactical Fighter Weapons Center

TO - Technical Order

TSPI - Time Space Position Information

TWX - Telegraphic Message

UDS - Universal Documentation System

2. TERMINOLOGY

ABORT. Failure to accomplish a mission for any reason other than enemy action. It may occur at any point from initiation of operation to destination.

ACCESSORY. A part, subassembly or assembly designed for use in conjunction with or to supplement another assembly, or a unit or set, contributing to the effectiveness thereof without extending or varying the basic function of the assembly or set. An accessory may be used for testing, adjusting, or calibrating purposes. (Examples: Instrument test, recording camera for radar set, head phones, emergency power supply.)

ACCURACY. The closeness of agreement between an experimentally obtained value and a true value.

ACQUISITION LIFE CYCLE. A series of phases through which a system or equipment passes during its lifetime. For major defense systems, the acquisition life cycle normally consists of five phases identified as Conceptual, Validation, Full-Scale Development, Production and Deployment with decision points between each of the first four phases. The decision points are Program Decision, Ratification Decision, and Production Decision. For certain nonmajor defense systems, the acquisition life cycle may be modified or abbreviated with regard to either the Conceptual, Validation, or Full-Scale Development Phases with decision points as prescribed by management to satisfy the needs for management visibility, review and control.

ADVANCED DEVELOPMENT. A term used in the research and development (R&D) of a system or equipment to denote that effort normally occurring after Research and Exploratory Development and prior to Engineering and Operational Deployment. It includes all R&D projects which have moved into the development of hardware for experimental or operational test. With relation to the acquisition life cycle, advanced development generally occurs at some point in the cycle between late Conceptual Phase and early Full-Scale Development Phase.

ADVANCED DEVELOPMENT MODEL. An item used for experimentation or tests to (a) demonstrate the technical feasibility of a design, (b) determine its ability to meet existing performance requirements, (c) secure engineering data for use in further development and, where appropriate, (d) establish the technical requirements for contract definition. Depending upon the complexity of the equipment and the technological factors involved, it may be necessary to produce several successive models to achieve additional objectives. The final advanced development model approaches the required form factor and employs standard parts (or nonstandard parts approved by the agency concerned). Serious consideration is given to military requirements such as reliability, maintainability, human factors and environmental conditions.

ASSEMBLY/SUBASSEMBLY.

ASSEMBLY. A number of parts or subassemblies or any combination thereof joined together to perform a specific function. (Examples: Power shovel front fork assembly, audio-frequency amplifier).

SUBASSEMBLY. Two or more parts which form a portion of an assembly or a unit replaceable as a whole, but having a part or parts which are individually replaceable. (Example: Gun mount stand, window recoil mechanism, floating piston, telephone dial, IF strip mounting board with mounted parts, power shovel dipper stick).

ATTACHMENT. A part, subassembly or assembly designed for use in conjunction with another assembly or a unit or set, contributing to the effectiveness thereof by extending or varying the basic function of the assembly, unit or set. (Examples: Hoisting attachment on a truck, milling attachment for a lathe).

AVAILABILITY. The probability that an item is in the operable and committable state at the start of the mission, when the mission is called for at an unknown (random) point in time.

BACKGROUND FACTOR. An independent test variable not included as a primary factor.

BEST ESTIMATE. The estimate of population parameter derived by any one of several methods deemed most desirable by some unchanging criterion. There is no single universally-best estimate.

BIAS. An unvarying tendency to vary from the true or desired value in one direction.

BLOCK. A planned homogeneous group. A grouping of several trials with a single level of some background factor or combination of background factors for the purpose of maximizing homogeneity among those units.

BOMB. A bomb (weapon) is an air launched unguided explosive device without powered flight.

"SMART" BOMB. A "smart" bomb (weapon) is an air launched, guided, nonpowered explosive device.

BREADBOARD. An arrangement in which components are fastened temporarily to a board or chassis for experimental work.

CALL SET-UP TIME. The time required to establish a complete circuit from the sender to the receiver given the call is not blocked.

CAPABILITY. A measure of the ability of an item to achieve mission objectives given the conditions during the mission. It may be stated as the probability that an item will achieve the mission objectives, given the Dependability.

CIRCULAR ERROR PROBABLE (CEP). The radius of a circle (centered at the expected arithmetic mean point of impact) within which half the missiles or projectiles are expected to fall.

COMMAND AND CONTROL SYSTEM. The facilities, equipment, communication, procedures, and personnel essential to a commander for planning, directing, and controlling operations of assigned forces pursuant to the missions assigned.

COMMUNICATIONS-ELECTRONICS. The broad field of activity encompassing the functions of program formulation, policy planning, inspection, and direction of communications electronics operation and maintenance. It includes supervisory and technical responsibilities for the construction, installation, operation, and maintenance of communications and electronics systems and equipment. It further includes all radio, wire, and other means used for the electrical and visual transmission and reception of information or messages in the clear or by cryptographic means; all radar and radiation aids to air traffic control and navigation and enemy aircraft warning and interception; all ground electronic devices and systems for the control and tracking of aircraft and guided missiles, electronic weather equipment, electronic countermeasures devices, and related electronic systems and equipment.

COMMUNICATIONS INTELLIGENCE. The technical and intelligence information derived from foreign communications by other than the intended recipients. Also called COMINT.

COMMUNICATIONS SECURITY. The protection resulting from all measures designed to deny unauthorized persons information of value which might be derived from the possession and study of telecommunications, or to mislead unauthorized persons in their interpretation of the results of such possession and study. COMSEC includes (1) cryptosecurity; (2) transmission security; (3) emission security; and (4) physical security of communications security materials and information.

a. (cryptosecurity) - The component of communications security which results from the provision of technically sound cryptosystems and their proper use.

b. (transmission security) - The component of communications security which results from all measures designed to protect transmissions from interception and exploitation by means other than cryptoanalysis.

c. (emission security) - The component of communications security which results from all measures taken to deny unauthorized persons information of value which might be derived from intercept and analysis of compromising emanations from cryptoequipment and telecommunications systems.

d. (physical security) - The component of communications security which results from all physical measures necessary to safeguard classified equipment, material, and documents from access thereto or observation thereof by unauthorized persons.

COMPATIBILITY. See "Operational Compatibility".

CONCEPTUAL PHASE. The initial period when the technical, military, and economic bases for acquisition programs are established through comprehensive studies and experimental hardware development and evaluation. The outputs are alternative concepts and their characteristics (estimated operational, schedule, procurement, costs and support parameters) which serve as inputs to the Development Concept Paper

(DCP) on major systems, Program Memoranda (PM) on smaller systems/equipment, and to HQ USAF decision documents (Program Management Directives) for programs that do not require OSD decisions.

CONFIDENCE LIMIT. The bounds on an interval estimate of a parameter.

CONFIDENCE REGION. The area or volume enclosed by a multidimensional interval estimate, whether of similar parameters on different axes or of different parameters.

CONFIGURATION. The functional and/or physical characteristics of hardware/software as set forth in technical documentation and achieved in a product.

CONFOUNDING. A mixing of effects such that results cannot be attributed to a particular single variable or group of variables.

CONTRACT. The legal agreement between DOD and industry, or similar internal agreement wholly within the government, for the development, production, maintenance or modification of an item(s).

DATA. The generic term "data" refers to recorded information that describes or refers to objects, conditions, ideas, situations, or other factors. It typically includes test measurements or observations and management, technical, and logistics information and reports. Broadly characterized, data may be objective or subjective. Objective data is observable, verifiable data that reflects actual conditions and is independent of the judgment of the individual recording the information. Subjective data is data that is dependent upon the judgment of the data collector and which reflects his unique frame of reference. Representative data includes:

(1) Administrative reports. Reports which require financial information of any sort, or contract production progress, socio-economic data, cost information, etc.

(2) Technical reports. Any technical document written to permanently record technical information, conclusions, and recommendations developed on scientific, technical, and engineering activities relating to a single task, project or contract, or a small group of closely related efforts. A technical report may be definitive for the subject presented, exploratory in nature, or a record of inconclusive or negative findings.

(3) Other data. Data needed by the Air Force to develop, acquire, install, test, operate, maintain, overhaul, repair, modify, supply, support, and reprocure systems and equipment. Such data may appear in the form of reports, technical manuals, charts, photographs, films, lists, tapes, drawings, specifications, parts breakdowns, etc.

RAW DATA. The original form of data at the time it was recorded.

DATA COLLECTION. The process of capturing and recording raw data, to include the gathering of the recorded data to a prescribed location.

DATA ELEMENT. The most basic piece of information. Usually a specific measurement or observation with its own unique descriptive name and physical condition.

DATA ITEM. A physical recording of a unique collection of data elements.

DATA MANAGEMENT. The process of determining and validating each data requirement and of planning for the timely and economical acquisition of data.

DATA REDUCTION. The process of transforming raw data into useful, ordered, or simplified form.

DATA VERIFICATION. The process of assessing whether data correctly represents the variable it characterizes and insuring that sufficient data is collected to support a test design.

DEBUGGING. A process to detect and remedy inadequacies, preferably prior to operational use.

DEFENSE SYSTEMS ACQUISITION REVIEW COUNCIL. (DSARC) An advisory council established by and functioning for the Secretary of Defense (SECDEF) to appraise SECDEF of the status and readiness of each major defense system to proceed to the next phase of effort in its acquisition life cycle. The permanent membership of the Council consists of the Assistant Secretary of Defense (Installations and Logistics), the Assistant Secretary of Defense (Comptroller), the Assistant Secretary of Defense (Program Analysis & Evaluation) and the Director of Defense Research and Engineering. The function of the Council is to evaluate the status of each candidate major defense system at three basic milestone points in the system acquisition process and to advise SECDEF concerning decisions required at each milestone. The milestone points (SECDEF decision required) are:

a. When the Sponsoring Service desires to initiate contract definition or equivalent effort. (Program Decision).

b. When it is desired to proceed from contract definition to Full-Scale development. (Ratification Decision).

c. When it is desired to transition from Development to Production for service Deployment. (Production Decision).

During DSARC reviews conducted for milestone points a. and b. above, the DSARC chairman is DDR&E. For the DSARC review identified in c. above, the chairman is ASD (I&L).

DELTA (GREEK LETTER Δ). A term or symbol used to denote a change or difference in a quantity.

DEPENDABILITY. A measure of the item operating condition at one or more points during the mission, including the effects of Reliability, Maintainability and Serviceability, given the item condition(s) at the start of the mission. It may be stated as the probability that an item will (a) enter or occupy any one of its required operational modes during a specified mission, (b) perform the functions associated with those operational modes.

DEPLOYMENT PHASE. The period beginning with the user's acceptance of the first operational unit and extending until the system is phased out of the inventory. It overlaps the Production phase.

DESCRIPTOR (STATISTICAL). A statistic or a population parameter.

DEVELOPMENT CONCEPT PAPER (DCP). The DCP is a coordinated management document which serves as:

a. The vehicle for major program decisions by the Secretary of Defense.
b. The record of basic program information, decision rationale, and review thresholds.
c. The instrument to effect implementation of these decisions. When approved by the Secretary of Defense, the DCP serves as authority to proceed with a particular phase of the acquisition cycle. Normally, the Air Force prepares the draft DCP from a DCP outline proposed (approved) by the OSD.

DEVELOPMENT TEST AND EVALUATION (DT&E). Test and evaluation which focuses on the technological and engineering aspects of the system, subsystem, or equipment items.

DRAFT TEST DIRECTIVE (TD). The document produced by AFTEC from the Initial Test Directive or PMD. The draft TD incorporates required testing and specifies required resources. It is submitted to HQ USAF for review, approval, and if appropriate, publication and dissemination as the HQ USAF Test Directive. See: Test Directive.

DSARC BRIEFING. A briefing on IOT&E projects to aid the Defense System Acquisition Review Council (DSARC) in making production decisions on major programs. The production decision is the last milestone prior to the Production and Deployment Phases in the acquisition life cycle. This decision determines whether to produce the system for operational use, defines the initial quantity to be procured, and approves plans for future production.

ENVIRONMENT. The aggregate of all conditions and influences which affect the operation of an item; e.g., physical location, temperature, humidity, pressure, shock, etc.

ELECTRONIC COUNTER-COUNTERMEASURES. The division of electronic warfare involving actions taken to insure friendly effective use of the electro-magnetic spectrum despite the enemy's use of electronic warfare. See also electronic countermeasures.

ELECTRONIC COUNTERMEASURES. The division of electronic warfare involving actions taken to prevent or reduce an enemy's effective use of the electromagnetic spectrum. It includes:

1. Jamming. The deliberate radiation, reradiation, or reflection of electromagnetic energy with the object of impairing the use of electronic devices, equipment, or systems being used by an enemy.

2. Deception. The deliberate radiation, reradiation, alteration, absorption, reflection of electromagnetic energy in a manner intended to mislead an enemy in the interpretation or use of information received by his electronic systems. It includes:

(a) Imitative Deception. Introducing radiations into enemy channels which imitate his own emissions.

(b) Manipulative Deception. The alteration or simulation of friendly electromagnetic radiations to accomplish deception. See also electronic counter-countermeasures.

ELECTRONIC INTELLIGENCE. The intelligence information product of activities engaged in the collection and processing, for subsequent intelligence purposes, of foreign, noncommunications, electromagnetic radiations emanating from other than nuclear detonations and radioactive sources. Also called ELINT.

EMISSION SECURITY. That component of communications security which results from all measures taken to deny unauthorized persons information of value which might be derived from intercept and analysis of compromising emanations from crypto-equipment and telecommunications systems.

ENGINEERING CHANGE. An alteration in the configuration of a configuration item or item, delivered, to be delivered, or under development, after formal establishment of its configuration identification.

ENGINEERING DEVELOPMENT MODEL (SERVICE TEST). An item used in tests to determine tactical suitability for military use in real or simulated environments for which the item was designed. It closely approximates an initial production design, has the required form, employs standard parts (or nonstandard parts approved by the agency concerned) and meets the standard military requirements such as reliability, maintainability, human factors, environmental conditions, etc.

ESTIMATE. An approximation of the value of a population parameter, based on a statistic calculated from a sample of the same population.

EVALUATIONS. The review and analysis of qualitative data produced during current or previous testing or operational usage, or combinations thereof to determine the worth of the item tested.

EXPLORATORY DEVELOPMENT MODEL. An item (preliminary parts or circuits) used for experimentation or tests to investigate or evaluate the feasibility and practicality of a concept, device, circuits, or system in breadboard or rough experimental form, without regard to the eventual overall fit or final form.

FACILITY. Any fixed installation which is an intimate part of a system. This includes real property installed equipment (RPIE).

FACTORIAL ARRANGEMENT. A design for primary factors in which a single experiment is used to investigate two or more levels of each of two or more factors.

FACTOR (STATISTICAL DEFINITION). An independent test variable.

FAILURE. The inability of an item to perform within previously specified limits.

FIDELITY (COMMUNICATIONS/ELECTRONICS). The exactness with which the information in the output of the Receiving Subsystem represents the information in the input to the transmitting subsystem.

FIGURE OF MERIT. A measure of effectiveness through which quantitative system requirements and characteristics can be related to mission objectives in optimizing the system design.

FOLLOW-ON OT&E (FOT&E). That OT&E accomplished on production items.

FORCE. A force is an aggregation of military personnel, weapon systems, vehicles and necessary support, or combination thereof.

FULL-SCALE DEVELOPMENT PHASE. The period when the system/equipment and the principal items necessary for its support are designed, fabricated, tested, and evaluated. The intended output is, as a minimum, a pre-production system which closely approximates the final product, the documentation necessary to enter the Production Phase, and the test results which demonstrate that the production product will meet stated requirements.

FUNCTION (PERFORMANCE). A product specification which states (1) the complete performance requirements of the product for the intended use, and (2) the necessary interface and interchangeability characteristics. It covers form, fit and function.

FUNCTIONAL CHARACTERISTICS. Quantitative performance, operating and logistic parameters and their respective tolerances. Functional characteristics include all performance parameters, such as range, speed, lethality, reliability, maintainability, safety.

GROUP. A collection of units, assemblies, or subassemblies which is not capable of performing a complete operational function. A group may be a subdivision of a set or may be designed to be added to or used in conjunction with a set to extend the function or the utility of the set. (Example: Antenna group).

HANDS-ON OPERATION. An operation involving the actual hardware of a system; contrasted with simulated operation.

HANDS-ON-TESTING. The active participation of personnel in the test being performed in roles which are essential to the conduct of the test.

HARMONIZATION. Requirements coordination activity between HQ USAF, the other Services, other departments of Government, and our allies.

HARDWARE/SOFTWARE. Hardware or software, or a combination of both, in which the software includes only that associated with hardware for operational use, e.g., computer programs for command and control, handbooks for operations, maintenance, etc., and excludes fabrication specifications, drawings, etc.

HUMAN FACTORS. Human factors are human psychological characteristics relative to complex systems, and the development and application of principles and procedures for accomplishing optimum man-machine integration and utilization. The term is used in a broad sense to cover all biomedical and psychosocial considerations pertaining to man in the system.

INHOMOGENEITY. Lack of eveness of nature or composition.

INITIAL OT&E. That OT&E which is initiated early in the development testing cycle and which primarily supports the first major production decision. For some systems, IOT&E may be continued after the first major production decision.

INTEGRATED LOGISTIC SUPPORT. A composite of the elements necessary to assure the effective and economical support of a system or equipment at all levels of maintenance for its programmed life cycle. The elements include all resources necessary to maintain and operate an equipment or weapons system, and are categorized as follows: (1) planned maintenance, (2) logistic support personnel, (3) technical logistic data and information, (4) support equipment, (5) spares and repair parts, (6) facilities, and (7) contract maintenance.

INITIAL TEST DIRECTIVE. A HQ USAF directive that initiates for planning purposes, an OT&E project. Sufficient authority and information is conveyed in this interim directive to form an initial test team and to develop the test design and test plan. The Initial Test Directive may be used in lieu of or in conjunction with a PMD to initiate test planning.

INTEROPERABILITY. The ability of systems, units or forces to provide services to and accept services from other systems, units or forces and to use the services so exchanged to enable them to operate effectively together.

ITEM. Any level of hardware assembly (i.e., system, segment of a system, subsystem equipment, component, part., etc.)

ITEM LEVELS. Item levels from the simplest division to the more complex are as follows:

Part
Subassembly
Assembly
Unit
Group
Set
Sub-system
System

JOINT TEST AND EVALUATION. An operational test and evaluation directed by the OSD in which two or more military components participate jointly in the preparation of the OT&E test plan, the conduct of the OT&E, and the preparation of the OT&E final report.

LETHALITY. The probability that a weapon will damage the military objective to a specified degree.

LEVEL. An alternate setting of a variable, whether quantitative or qualitative.

LIFE CYCLE COSTS. Life cycle costs are the costs of acquisition plus operation and logistic support costs for the specified operational lifetime.

LOGISTICS SUPPORTABILITY. How well the composite of support considerations necessary to achieve the effective and economical support of a system or equipment for its life cycle meets stated quantitative and qualitative requirements.

MAINTAINABILITY. A characteristic of design and installation which is expressed as the probability that an item will conform to specified conditions within a given period of time when maintenance action is performed in accordance with prescribed procedures and resources.

MAJOR PROGRAM. A major program is a system acquisition program designated by the Secretary of Defense having:

a. an estimated RDT&E cost in excess of 50 million, or
b. an estimated production cost in excess of 200 million, or
c. National urgency, or
d. been recommended by DOD Component Heads or Office of Secretary of Defense (OSD) officials.

MANEUVERABILITY. The ability to perform a change, or combination of changes, in altitude, airspeed, and direction.

MEAN TIME-BETWEEN FAILURES (MTBF). For a particular interval the total functioning life of a population of an item divided by the total numbers of failures within the population during the measurement interval.

MEAN TIME-TO-REPAIR (MTTR). The total corrective maintenance time divided by the total number of corrective maintenance actions during a given period of time. The corrective maintenance time includes preparation, fault location, item obtained, fault corrections adjustment-calibration, check-out- and clean-up times.

MEASURE OF EFFECTIVENESS (MOE). A parameter which evaluates the extent of the adequacy of the item to accomplish an intended mission under specific conditions.

MEASURES OF PERFORMANCE. The measures Availability, Dependability and Capability.

MEASUREMENT ERROR. An error due to the imperfect accuracy and/or precision of the measuring process.

MILITARY UTILITY. The military/operational value of an item/system when measured from within a current concept of operation.

MISS DISTANCE. A radial measure of nearness to a target.

MISSILE. A missile weapon is a guided and powered explosive device.

MISSILE CLUSTER. A group of one or more missiles of the same type that are loaded onto a launcher, following which the missiles and launcher are mated as a single unit to an aircraft.

MISSION. 1. A mission is a task, together with the purpose, which clearly indicates the action to be taken and the reason therefor. 2. In common usage, especially when applied to lower military units, a duty assigned to an individual or unit, a task. 3. The dispatching of one or more aircraft to accomplish one particular task.

MOCK-UP. An assembly having special contractual or engineering significance with relation to a system or subsystem, but which is not required solely for the conduct of either development tests, technical evaluation, or operational evaluation.

MODEL. Any device, technique, or process by means of which the specific relationships of a set of quantifiable system parameters may be investigated.

MODEL. Any device, technique, or process by means of which the specific relationships of a set of quantifiable system parameters may be investigated.

MODEL LEVELS (DEVELOPMENT & PRODUCTION) (Lowest to highest)

Exploratory development model

Advanced development model

Engineering development (service test) model

Preproduction (prototype) model

Production model

MULTISYSTEM - See "System."

NEAR REAL TIME. Delay caused by automated processing and display, between the occurrence of an event and the reception of the data at some other location.

OPERATING COMMAND. The command primarily responsible for the operational employment of a system, subsystem, or item of equipment. This term generally applies to those operational commands or organizations designated BY HQ USAF to conduct or participate in operations or operational testing.

OPERATIONAL CAPABILITY. A measure of the ability of an item to achieve mission objectives given the conditions during the mission.

OPERATIONAL COMPATIBILITY. The capability of two or more operational items/systems to exist or function as elements of a larger operational system or operational environment without mutual interference.

OPERATIONAL EFFECTIVENESS. An assessment of the effectiveness of a system in meeting its defined operational requirements in its intended operating environment, including effects of countermeasures and tactics.

OPERATIONAL RELIABILITY. (See reliability).

OPERATIONAL REQUIREMENT. The need of a military capability essential to the accomplishment of approved objectives, missions and tasks.

OPERATIONAL SUITABILITY. The measure of how well the system is suited to be operated and maintained by military personnel in the field. It involves quantitative and qualitative assessments of reliability, maintainability, supportability, and operability,

OPERATIONAL SYSTEMS DEVELOPMENT. The activity which includes a research and development effort directed toward development, engineering and test of systems, support programs, vehicles and weapons that have been approved for production and Service employment.

OPERATIONAL TEST AND EVALUATION (OT&E). Subsystem, system or multisystem test and evaluation conducted under Service operating conditions, insofar as practicable, to determine and/or evaluate the operating characteristics of that (specified) subsystem, system, or multisystem.

OT&E FINAL REPORT. A formal document prepared by or under the guidance of the Test Director, and published and promulgated after the completion of a specific assignment of OT&E, which provides background information, test item description, test procedures, results, conclusions, and recommendations.

OT&E STAFF OFFICER. HQ level OPR and focal point for an assigned OT&E project.

OT&E TEAM. A team formed by the AFTEC Commander to conduct specific OT&E projects. The OT&E team consists of organically assigned AFTEC personnel and other resources (personnel, forces, equipment, etc.) provided by Major Commands (MAJCOMs) in accordance with a HQ USAF PMD or Test Directive (TD).

OT&E TEST DIRECTOR. A person designated by the AFTEC Commander to conduct a specific test and evaluation project. Test directors are designated from AFTEC personnel or, if appropriate, and with the MAJCOM Commander's concurrence, from MAJCOMs personnel assets.

OT&E TEST OFFICER. An officer supporting the OT&E Test Director in planning and conduct of the test.

OT&E TEST PLAN. A formal document prepared by or under the direction of the assigned Test Director which provides the complete detailed, coordinated and integrated plan for the time-phased tasks and resources required to conduct, analyze, and report an OT&E task assigned by an OT&E Test Directive.

PARAMETER. A measurable characteristic of a population.

PARTICIPATING COMMAND. Any USAF, DOD, Federal, or contractor organization conducting, supporting, or participating in testing.

PERFORMANCE. The technical, operational, and support characteristics of systems, subsystems, equipments, or modifications.

PERSONNEL ERROR. Incorrect performance of required duties by operating or maintenance personnel which cause a failure.

POINT ESTIMATE. A single-valued estimate of a parameter.

POPULATION. The set of all situations, items, or events from which a sample may be drawn and/or about which inferences may be made. Sometimes the target population is distinguished from the sample population; the former is the set of all situations, items, or events about which it is desired to draw inferences while the latter is the subset of the target population about which it is permissible to draw statistical inferences because it is the subset from which samples were legitimately drawn.

POSITION. The term used to denote the location, in three dimensions, of an item with the origin of the coordinates taken as the system under test. For systems that cover a significant area, the origin should be defined by the tester as some significant point within the confines of the system.

PRECISION. The spread of several experimentally-obtained values. The magnitude of the random errors in measuring or estimating a value.

PREPRODUCTION (PROTOTYPE) MODEL. An item suitable for complete evaluation of form, fit and performance. It is in final form in all respects, employs standard parts (or nonstandard parts approved by the agency concerned) and is completely representative of final equipment.

PRIMARY FACTOR. An independent variable that the test is designed to answer questions about. Some means of control in operational employment is implied.

PROBABILITY OF VOICE INTELLIGIBILITY. The probability of sentence (or word) recognition as determined by test listeners when standard test is spoken by test talkers.

PROBABILITY OF VOICE NATURALNESS. The probability that the listener recognizes the talker's voice.

PRODUCTION MODEL (ITEM). An item in its final form of final production design made by production tools, jigs, fixtures and methods. It employs standard parts (or nonstandard parts approved by the agency concerned).

PRODUCTION PHASE. The period within the acquisition life cycle from production approval until the last system/equipment is delivered and accepted. The objective is to efficiently produce and deliver effective and supportable systems to the operating units. It includes the production and deployment of all principal and support equipment.

PROGRAM ADVOCACY. That effort expended during the Conceptual Phase to obtain program approval. The extent of this effort varies with each proposal, is specified by HQ USAF, and will be the minimum required by the approving authority. For major programs, it supports a request to the Secretary of Defense to proceed with the Validation Phase and contains the basic information from which the Development Concept Paper (DCP) is prepared. On other programs, it supports the preparation of the program/project approval document.

PROGRAM MANAGEMENT DIRECTIVE (PMD). The official HQ USAF management directive used during the entire system acquisition cycle to provide direction to the implementing and participating commands and to serve as a contract between those parties. The content of the PMD, including the required HQ USAF review and approval actions, is tailored to the needs of each individual acquisition program; however, it normally serves to (a) Define the authority of the Program Manager. (b) State the responsibilities of participating commands. (c) State resource requirements. (d) Request studies and analyses. (e) Initiate, approve, change, modify or terminate system acquisition program actions. (f) Satisfy document needs for program advocacy, development, production and modification funded by RDT&E or procurement funds. (g) Define test objectives (when the Test Objectives Annex (TOA) is included).

PROGRAM MANAGEMENT PLAN (PMP). The document developed and issued by the Program Manager which shows the integrated time-phased tasks and resources required to complete the task specified in the PMD. The PMP is tailored to the needs of each individual program.

PROGRAM MANAGER. The single Air Force manager during any specific phase of the acquisition life cycle (System Program Director, Program Manager, or System Manager/ Item Manager).

PROGRAM MEMORANDUM (PM). An OSD document prepared with similar format, content and coordination as the DCP but documents program guidelines, and thresholds for those significant development programs which are not subject to specific DCP action.

PROGRAM OFFICE. The field office organized by the Program Manager to assist him in accomplishing the program tasks.

PROPAGATION MEDIUM SUBSYSTEM (COMMUNICATIONS ELECTRONICS). That subsystem used to connect the output of the transmitting subsystem to the input of the receiving subsystem. It includes space, earth, water, coax, field wire, waveguides, light pipes, etc.

PROTOTYPE. The first full scale functional form of a new system, subsystem or component on which the design of subsequent production items is patterned.

RANDOM ERROR. A completely unpredictable (magnitude and direction) error.

RANDOMIZATION. An intentionally-unpatterned association of variables and levels of variables such that the effect of one (level or variable) is not confounded with the effect of another (level or variable).

RAW DATA. (See "Data".)

REAL TIME. Having no delay, except for the time required for the transmission by electromagnetic energy, between the occurrence of an event or the transmission of data, and the knowledge of the event, or reception of the data at some other location.

RECEIVING SUBSYSTEM (COMMUNICATIONS/ELECTRONICS). The subsystem that takes the output from the propagation medium subsystem, processes it, and returns it to useful information. It includes the input networks (antenna, antenna drive, input matching network, etc.) tuning, amplifying, demodulating, decrypting, output, equipment, (speakers, printers, visual display, etc), and operators.

REGRESSION. The locus of mean values of a variable for given values of other variables.

RELIABILITY. The measure of probability that an item will perform its intended function for a specified interval under stated condition." The probability that a system, subsystem, or equipment will perform a required function under specified conditions, without failure, for a specified period of time.

REPLICATION. A repetition of some combination of primary factors on some combination of background factors or some group of combinations of primary factors on some group of combinations of background factors. An attempt to run multiple trials under the same conditions.

REPORTING TIME INTERVAL. (1) In surveillance, the time interval between the detection of an event and the receipt of a report by the user. (2) In communications, the time for transmission of data or a report from the originating terminal to the end receiver.

RESOURCES. The requirements of an organization for the conduct of assigned tests. It includes funds, manpower, physical plant (equipment and real property), test instrumentation equipment, standard vehicles, aircraft communications and electronics, data reduction, and associated software, including computer programs.

RISK. 1. (System acquisition). The technological and operational uncertainty, present in progressive stages of system acquisition, of the success of the development, production, and deployment of the system and of the capability of the system to satisfy design objectives.

2. α Risk - (test design). The probability of rejecting a true hypothesis.

3. β Risk - (test design). The probability of accepting a false hypothesis.

ROCKET. A rocket weapon is an unguided explosive device with powered flight.

SAMPLE. The situations, items, or events which are actually observed in an experiment.

SPHERICAL ERROR PROBABLE (SEP). The radius of a sphere (centered at the expected arithmetic mean point of closest arrival) within which half the missiles or projectiles are expected to pass.

SQUARE. A simultaneous grouping of several trials into orthogonal sets of blocks defined by the levels of two or more variables.

STANDARD. A document created and promulgated to control variety.

STATISTIC. A measurable characteristic of a sample (or samples).

SUBSYSTEM - (See "System.")

SCENARIO. A description of a segment of a major mission category and level of activity within that category which shows the interaction of equipment, personnel, and procedures. A scenario includes equipment and facilities specifically utilized and personnel specifically required in the segment, and provides general procedures relating equipment and personnel.

SET. A unit or units and necessary assemblies, subassemblies and parts connected together or used in association to perform an operational function. (Examples: Radio receiving set, radar homing set). "Set" is also used to denote a collection of related items such as a 'tool-set,' drawing set,' or a 'set of tires'.

SIGNAL INTELLIGENCE. A generic term which includes both communication intelligence and electronic intelligence. Also called SIGINT.

SIGNAL-TO-NOISE-RATIO. The ratio of signal power to noise power at a selected point in a system (See Appendix G).

SORTIE. An operational flight by one aircraft.

SPEED OF SERVICE. (COMMUNICATIONS/ELECTRONICS). The time required for a message to move through the system from the first bit into the transmitting subsystem to the last bit out of the receiving subsystem or its equivalent.

SUPPORT EQUIPMENT. That equipment required to make an item, system, or facility operational in its intended environment. This includes (a) all equipment required to maintain and operate the item, system or facility including aerospace ground equipment and ground equipment and (b) computer programs related thereto.

SUPPORTING COMMAND. A command that provides direct support to a system or test program. Normally, the term refers to AFLC, USAFSS, and ATC in their role as logistics support and training organizations.

SURVIVABILITY. The capability of a system to withstand a man-made hostile environment without suffering an abortive impairment of its ability to accomplish its designated mission.

SYSTEM. A system is a composite of equipment, skills and techniques capable of performing and/or supporting an operational role. A complete system includes all equipment, related facilities, material, software, services, and personnel required for its

operation and support to the degree that it can be considered a self-sufficient unit in its intended operational environment.

a. Multisystem - a group of two or more systems used to perform and/or support a mission.

b. Subsystem - A composite of equipment, skills, and techniques which perform a unique function, but which is not self-sufficient to perform the complete mission.

SYSTEM DEFINITION. Determination of qualitative performance and physical requirements which are adequate for design of a system element.

SYSTEM EFFECTIVENESS. A measure of the degree to which a system achieves a set of specific mission requirements. It is a function of availability, dependability, and capability.

TACTICS. 1. The employment of units in combat. 2. The ordered arrangement and maneuver of units in relation to each other and/or to the enemy in order to utilize their full potentialities.

TELECOMMUNICATIONS. Any transmission, emission, or reception of signs, signals, writing, images, and sounds or intelligence of any nature by wire, radio, visual, or other electromagnetic system.

TEST. A critical examination, observation, or evaluation.

TEST DESIGN. (1) That part of the plan for selecting a group of trials and making observations on those trials, such that the maximum amount of information can be derived with the minimum expenditure of resources. (2) Also, the activity which develops (1).

TEST DIRECTIVE. A HQ USAF directive which authorizes and directs the planning, execution, and reporting of a test program. It also tasks the MAJCOMs to provide the required resources for an AFTEC assigned project.

TEST PLAN. (See OT&E Test Plan).

TEST PROCEDURE. A document which describes the detailed requirements for conducting one type of test.

TEST PROFILE. A sequence of time phased actions describing the role of the test item during an individual test.

TEST REPORT. (See OT&E Final Report.)

TIME ACTIVE. That time during which an item is in the operational inventory.

TIME, ADJUSTMENT OR CALIBRATION. That element of Maintenance Time during which the needed adjustment of calibrations are made.

TIME, ADMINISTRATIVE. Those elements of Delay Time that are not included in Supply Delay Time.

TIME, ALERT. That element of Uptime during which an item is thought to be in specified operating condition and is awaiting a command to perform its intended mission.

TIME, CHECKOUT. That element of Maintenance Time during which performance of an item is verified to be in specified condition.

TIME, CLEANUP. That element of Maintenance Time during which the item is enclosed and extraneous material not required for operation is removed.

TIME, DELAY. That element of Downtime during which no maintenance is being accomplished on the item because of either supply delay or administrative reasons.

TIME, DOWN (DOWNTIME). That element of Time during which the item is not in condition to perform its intended function.

TIME, FAULT CORRECTION. That element of Maintenance Time during which a failure is corrected by (a) repairing in place; (b) removing, repairing, and replacing; or (c) removing and replacing with a like serviceable item.

TIME, FAULT LOCATION. That element of Maintenance Time during which testing and analysis is performed on an item to isolate a failure.

TIME, INACTIVE. That time during which an item is in reserve (in the Inactive Inventory).

TIME, ITEM OBTAINMENT. That element of Maintenance Time during which the needed items or items are being obtained from designated organizational stockrooms.

TIME, MISSION. That element of Uptime during which the item is performing its designated mission.

TIME, MODIFICATION. The time necessary to introduce any specific change(s) to an item to improve its characteristics or to add new ones.

TIME, PREPARATION. That element of Maintenance Time needed to obtain the necessary test equipment and maintenance manuals, and set up the necessary equipment to initiate fault location.

TIME, REACTION. That element of Uptime needed to initiate a mission, measured from the time the command is received.

TIME, SUPPLY DELAY. That element of Delay Time during which a needed item is being obtained from other than the designated organizational stockrooms.

TIME, TURN-AROUND. That element of Maintenance Time needed to service or checkout an item for recommitment.

TIME, UP (UPTIME). That element of Active Time during which an item is either alert, reacting, or performing a mission.

TOLERANCE LIMITS. The bounds on an interval containing a specified fraction of a population of values.

TRANSMISSION SECURITY. The component of communications security that results from all measures designed to protect transmission from interception and exploitation by means other than cryptoanalysis.

TRANSMITTING SUBSYSTEM. That subsystem that takes the information to be sent, processes it and inputs it into the propagation medium. It therefore includes its power supply, operators, transmitter input (microphone, key, sensors, A to D converter, etc.)

encrypting, modulating, oscillator, power amplifier and feed network (antenna, antenna drive, output matching networks, etc.). It includes the power supply of transceivers.

TRIAL. A single unique opportunity for observing the value of a dependent variable.

TWO-SIDED TESTING. (1) (OT&E). Testing in which the presence and activity of an adversary is involved. (2) (Statistics). In hypothesis testing, a test for which both upper and lower confidence limits are finite.

VALIDATION PHASE. The period when major program characteristics are refined through extensive study and analyses, hardware development, test, and evaluations. The objective is to validate the choice of alternatives and to provide the basis for determinating whether or not to proceed into Full-Scale Development.

VARIABLE. Something that is able or apt to change.

VOICE INTELLIGIBILITY. The percent of meaningful records, phrases or sentences spoken by a talker (or talkers) that are heard correctly by a listener (or listeners).

VULNERABILITY. The characteristics of a system which cause it to suffer a definite degradation (incapability to perform the designated mission) as a result of having been subjected to a certain level of effects in unnatural (man-made) hostile environments.

Appendix I

REFERENCES

DOCUMENT NUMBERS	TITLE
	DEPARTMENT OF DEFENSE
DODD-3200.11	Use, Management & Operation of National Ranges
-4120-3-M	Standardization Policies & Procedures
-5000.1	Acquisition of Major Defense Systems
-5000.2	Decision Coordinating Papers (DCP) and the Defense Systems Acquisition Review Council (DSARC)
-5000.3	Test and Evaluation
-5000.9	Standardization of Military Terminology
-5010.19	Configuration Management
	JOINT CHIEFS OF STAFF
JCS PUB 1	Dictionary of Military & Associated Terms
	MILITARY STANDARDS
MIL-STD-499	System Engineering Management
MIL-STD-721B	Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors & Safety
MIL-STD-847A	Format Requirements for Scientific & Technical Reports
MIL-STD-1472	Human Engineering Design Criteria for Military Systems, Equipment & Facilities
	USAF
AF-HOI-11-16	Responsibilities, Functions & Procedures Pertaining to Development Concept Papers (DCPs)
AF-HOI-11-21	Responsibilities, Functions, & Procedures Pertaining to Area Coordinating Papers (ACPs)
AF-HP-20-1	The Organization, Doctrine & Procedural Concepts of the Air Staff
AF-HOI-21-18	The Air Force Board Structure

DOCUMENT NUMBERS	TITLE
	USAF (continued)
AF-HOI-27-1	DOD Programming System
AF-HP-80-10	Air Staff R&D System
AF-HOI-310-1	HQ USAF Program for the Management of Contractor Data
AF-HOI-375-2	Selected Acquisition Reports
AF-HOI-800-1	DCP/DSARC Preparation
AF-HOI-800-2	Program Management Direction
AF/RD-OI-12-1	AF/RD Documentation Management
AF/RD-OI-20-4	Research, Development & Acquisition Program Review Group
AF/XOOW-DDOI-10-2	Test Directive Administrative Practices
AFR 8-2	USAF Technical Order System
AFM 11-1	USAF Glossary of Standardized Terms
AFM 11-2	USAF Manual of Abbreviations
AFR 11-4	Host-Tenant Support Responsibilities of USAF Organizations
AFR 12-40	Documentation Storage & Retrieval
AFR 12-41	Engineering Data Service Centers
AFR 19-1	Protection & Enhancement of Environmental Quality
AFR 19-2	Environmental Assessments & Statements
AFM 50-18	Weapons Ranges
AFR 55-11	Programming of Requirements & Reporting Expenditures for Missile/Targets in Non-Combat Firing Programs
AFR 55-89	Tactical Fighter Weapons Delivery Qualification
AFR 57-1	Policies, Responsibilities & Procedures for Obtaining New & Improved Operational Capabilities
AFR 57-4	Retrofit Configuration Changes
AFR 57-5	Quick Reaction Capability
AFM 66-1	Maintenance Management
AFR 66-8	Maintenance Evaluation Program
AFR 66-14	Equipment Maintenance Policies, Objectives & Responsibilities

DOCUMENT NUMBERS	TITLE
	USAF (continued)
AFR 66-38	Nondestructive Inspection Program
AFR 66-44	Equipment Maintenance Quality & Reliability Assurance Program
AFM 66-267	Maintenance Data Collection System
AFR 67-19	Logistic Support of Research, Development, Test & Evaluation Activities
AFR 73-1	Defense Standardization Program
AFR 80-5	Reliability & Maintainability Programs for Systems, Subsystems, Equipment, & Munitions
AFR 80-11	Importance Categories
AFR 80-14	USAF Test & Evaluation
AFR 80-16	Characteristics Guides
AFR 80-18	DOD Engineering for Transportability
AFR 80-25	All-Weather Testing
AFR 80-38	Management of USAF Survivability Program
AFR 80-40	The Scientific & Technical Information Program
AFR 80-44	Defense Documentation Center for Scientific & Technical Information
AFR 80-45	Distribution Statements on Technical Documents
AFR 80-46	Management of Personnel Subsystem/Human Factors in System, Subsystem, Equipment & Modifications Program
AFR 100-4	Radio Frequency Management
AFR 100-11	Defense Communications Agency Operational Evaluation Program
AFR 102-5	USAF Management Policies Governing Development, Acquisition & Operation of Command Control Systems
AFM 126-1	Conservation & Management of Natural Resources
AFM 127-1	Aircraft Accident Prevention & Investigation
AFM 127-2	USAF Accident/Incident Reporting
AFM 127-100	Explosive Safety Manual

DOCUMENT NUMBERS	TITLE
	USAF (continued)
AFR 127-4	Investigating & Reporting USAF Accidents & Incidents
AFR 205-1	Information Security Program
AFP 205-2-1	Developing A Methodology for Security Classification of Scientific & Technical Material
AFR 205-29	Classification Criteria & Factors for Scientific & Technical Information
AFR 205-37	Security Classification Guides
AFR 300-5	Standardization of Data Elements & Related Features
AFM 400-25	Logistics Performance Measurement & Evaluation System
AFR 400-26	Logistics Support for Systems/Equipment Test Programs
AFR 800-2	Program Management
AFR 800-3	Engineering of Defense Systems
AFR 800-4	System/Equipment Turnover & Management Transition
AFR 800-5	Selected Acquisition Reports
AFR 800-8	Integrated Logistics Support (ILS) Program for Systems & Equipment
AFR 800-10	Management of Multi-Service Systems, Programs and Projects
	AFSC
AFSCP 80-3	USAF Technical Facility Capability Key
AFSCP 80-5	Guide for Design, Conduct, & Analysis of USAF Tests
AFSCR 80-10	Test Support for Systems & Equipment in the Deployment Phase
AFSC S1/AFR 80-14	AFSC Supplement to AFR 80-14, Test & Evaluation
AFSCR 80-24	AFSC Technical Facilities Register
AFSCM 400-25	Logistics Performance Measurement & Evaluation System
AFSCP 800-3	A Guide for Program Management
AFSCR 800-18	Joint Operational & Technical Review (JOTR)

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
RADC-TR-74-270, Volume I (of two)

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
STANDARD PROCEDURES FOR AIR FORCE OPERATIONAL TEST AND EVALUATION
Appendices

5. TYPE OF REPORT & PERIOD COVERED
Final Report

6. PERFORMING ORG. REPORT NUMBER
None

7. AUTHOR(s)
D.E. Simon (RCA)
et al

8. CONTRACT OR GRANT NUMBER(s)
F30602-73-C-0375

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Braddock, Dunn & McDonald, Incorporated*
5301 Central Avenue
Albuquerque, New Mexico 87108

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
PE 65804D
JO 31050102

11. CONTROLLING OFFICE NAME AND ADDRESS
Rome Air Development Center (IRAA)
Griffiss Air Force Base, New York 13441

12. REPORT DATE
October 1974

13. NUMBER OF PAGES
260

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
Same

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
N/A

16. DISTRIBUTION STATEMENT (of this Report)

Distribution limited to US Gov't agencies only; test and evaluation; October 74. Other requests for this document must be referred to RADC(IRAA),GAFB,NY 13441.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

*with subcontractors RCA Government and Commercial Systems Missile and Surface Radar Division and the Xerox Corporation Data System Division.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

OT&E
Initial OT&E
Weapon System Life Cycle
Measures of Effectiveness OT Planning
Joint OT&E Simulations
OT&E Data Handling
AF Testing
OT&E Mathematics and Statistics
OT&E Statistical Design
Constant Improvement Program
Test Report Formats

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes the overall structure of Air Force Operational Test and Evaluation process from the appearance of the test directive to the production of the final report. It describes the major documentation requirements applicable to major and minor weapon system acquisitions and provides guidance to the Air Force OT&E community in the areas of formulation of test objectives, selection of test concepts, determination of test planning criteria, OT&E data collection and analysis requirements, formulation of OT&E conclusions and

20. ABSTRACT (continued)

recommendations, and test reporting. It further describes procedures for the development of statistical design of an operational test. This report culminates an effort to standardize the management and analytical procedures applicable to Air Force Operational Test and Evaluation. Toward this end, standardized data elements of measures of effectiveness are developed.

MISSION

of

Rome Air Development Center

RADC is the principal AFSC organization charged with planning and executing the USAF exploratory and advanced development programs for electromagnetic intelligence techniques, reliability and compatibility techniques for electronic systems, electromagnetic transmission and reception, ground based surveillance, ground communications, information displays and information processing. This Center provides technical or management assistance in support of studies, analyses, development planning activities, acquisition, test, evaluation, modification, and operation of aerospace systems and related equipment.

Source AFSCR 23-50, 11 May 70